



Program to End Modern Slavery (PEMS)

Process Evaluation – Final Report

*U.S. Department of State, Office to Monitor and Combat Trafficking in Persons,
International Programs Section*

July 2021

This publication was produced at the request of the United States Department of State. It was prepared independently by EnCompass LLC.

PROGRAM TO END MODERN SLAVERY (PEMS)

Process Evaluation

Final Report

U.S. Department of State, Office to Monitor and Combat Trafficking in Persons, International Programs Section

July 29, 2021

IDIQ Contract Number: SAQMMA17A0367

Technical and Advisory Services for Program Evaluation Requirements

Task Order Number: 19AQMM19S0629

Contact:

Kimberly Norris, PhD
Senior Evaluation Specialist
Email: knorris@encompassworld.com

EnCompass LLC
1451 Rockville Pike, Suite 600
Rockville, MD 20852
Tel: +1-301-287-8700

DISCLAIMER

The authors' views expressed in this publication do not necessarily reflect the views of the United States Department of State or the United States Government.

CONTENTS

ACRONYMS.....	II
GLOSSARY.....	V
EXECUTIVE SUMMARY	VII
BACKGROUND	1
PEMS 1 Mandate.....	1
PEMS 1 Program	2
EVALUATION PURPOSE AND QUESTIONS	6
Evaluation Purpose	6
Reading this Report.....	6
Evaluation Questions	7
EVALUATION METHODS	10
Evaluation Approach and Timeline	10
Prevalence Estimation Data Collection Methods and Sampling	11
Project Implementation Data Collection Methods and Sampling	12
Data Analysis and Synthesis for Prevalence Estimation Studies and Project Implementation	13
Ethics, Data Security, Collaboration	14
Evaluation Challenges and Limitations.....	15
PEMFAC.....	18
Summary of Learning from the PEMFAC.....	19
PREVALENCE ESTIMATION STUDY ASSESSMENT.....	20
Summary of Prevalence Estimation Studies.....	20
PEMS 1 Prevalence Estimation Studies Assessment Methodology	21
Findings Across Prevalence Estimation Methodologies	23
Conclusions for Prevalence Estimation Studies.....	43
FINDINGS ACROSS PROJECT IMPLEMENTATION ASSESSMENT	49
Summary of Projects	49
Conclusions for PEMS Project Implementation.....	78
RECOMMENDATIONS.....	83
Recommendations for Funding and Conducting Future PE Studies Based on the Evaluation of PEMS 1 Prevalence Estimation Studies	83
Recommendations for Funding and Conducting Future Projects to Combat TIP Based on the Evaluation of PEMS 1 Implementation and Early Progress	85
CONSIDERATIONS FOR FURTHER WORK	89
Considerations for Funding Evaluation of Complex Efforts	89
Considerations for Utilizing TIP Definitions in Prevalence Estimation Studies	90
Considerations for Future Evaluation of Prevalence Estimation Studies	90
Considerations for Evaluating Projects and Subawardees	91
ANNEXES	92

ACRONYMS

BITC	Business in the Community
BOCW	Building and Other Construction Workers
CIFAL	International Training Center for Authorities and Leaders
CORE	Center for Reintegration and Empowerment
CSE	Commercial sexual exploitation
CSEC	Commercial sexual exploitation of children
CSO	Civil society organization
CWC	Child Welfare Committee
DAIS	Data Analysis, Integration, and Synthesis events
DOL	Department of Labor
DOLISA	Department of Labour, Invalids and Social Affairs (Vietnam)
DOS	Department of State
ECQ	Enhanced Community Quarantine
FEA	Fair Employment Association
FEF	Fair Employment Foundation
FGD	Focus group discussion
GAO	Government Accountability Office
GFEMS	Global Fund to End Modern Slavery
IACAT	Inter-Agency Council Against Trafficking
ICMS	Integrated case management system
IEC	Informal Education Communication
IFGS	Institute of Family and Gender Studies
ILO	International Labour Organization
IOM	International Organization for Migration
IRB	Institutional Review Board
IRIS	International Recruitment Integrity Standard

KII	Key informant interview
KPI	Key performance indicator
LMT	Longitudinal migration tracking
LOI	Line of inquiry
MEL	Monitoring, evaluation, and learning
METIP	Monitoring and Evaluation in Trafficking in Persons (Community of Practice)
NCR	(Delhi) National Capital Region
NDAA	National Defense Authorization Act
NGO	Nongovernmental organization
NOFO	Notice of Funding Opportunity
NRM	National Referral Mechanism
NSUM	Network scale-up method
ODI	Overseas Development Institute
OFW	Overseas Filipino worker
OWWA	Overseas Workers Welfare Administration
PDOs	Pre-departure orientation seminars
PE	Prevalence estimation
PEMFAC	Prevalence Estimation Methodology Features and Characteristics
PEMS	Program to End Modern Slavery
PPS	Probability proportional to size
RBA	Responsible Business Alliance
RCT	Randomized controlled trial
RDS	Respondent-driven sampling
RFP	Request for Proposal
RPEP	Research and Prevalence Estimation Panel
RPL	Recognition of Prior Learning
RRP	Responsible Recruitment Program
SDM	Sub-Divisional Magistrate

SHA	Sustainable Hospitality Alliance
SOW	Statement of work
STCI	Save the Children India
TESDA	Technical Education and Skills Development Authority
TIP	Office U.S. Department of State Office to Monitor and Combat Trafficking in Persons
TIP	Trafficking in persons
TLS	Time-location sampling
TVPA	Trafficking Victims Protection Act (2000)
UCLA	University of California, Los Angeles
UMass	University of Massachusetts
USAID	United States Agency for International Development
VAMAS	Vietnam Association of Manpower Supply
VASS	Vietnam Academy of Social Sciences
YCI	Youth Career Initiative

GLOSSARY

This glossary provides a list of concepts and terms relevant to the field of human trafficking and evaluation as used in the Evaluation Design for the Program to End Modern Slavery (PEMS) 1 process evaluation.

Delphi technique	A method of group decision-making and/or forecasting that involves successively collating the judgments of experts.
Forced labor	When a trafficker recruits, harbors, transports, provides, or obtains a person for labor or services by using force, fraud, or coercion (Source: Statement of Work). Also synonymous with trafficking in persons.
Logframe	A logframe graphically illustrates program components; creating it helps stakeholders clearly identify outcomes, outputs, inputs, and activities.
Modern slavery	“Modern slavery,” “trafficking in persons,” and “human trafficking” have been used as umbrella terms for the act of recruiting, harboring, transporting, providing, or obtaining a person for compelled labor or commercial sex acts through the use of force, fraud, or coercion. ¹
Prevalence	The proportion of a particular population affected by a specific condition at a given time.
Prevention	Anti-trafficking approaches that work to prevent trafficking by disseminating accurate and targeted information, strategically reaching at-risk populations to prevent recruitment, and building partnerships for awareness, expertise, and creative solutions (Source: Office to Monitor and Combat Trafficking in Persons [TIP Office] website).
Prosecution	Anti-trafficking approaches that help enable effective law enforcement action against traffickers, including criminalization of all forms of human trafficking, investigation and prosecution of human trafficking cases, and conviction of those responsible with sufficiently stringent sentences (Source: TIP Office website).
Protection	Anti-trafficking approaches that help identify victims, provide them with referrals for a comprehensive array of services, directly fund nongovernmental organizations to supply those services, and support these individuals as they rebuild their lives (Source: TIP Office website).

¹ The Trafficking Victims Protection Act (TVPA) of 2000 (Pub. L. 106-386), as amended, and the Protocol to Prevent, Suppress and Punish Trafficking in Persons, Especially Women and Children, supplementing the United Nations Convention against Transnational Organized Crime (the Palermo Protocol) describe this compelled service using a number of different terms, including involuntary servitude, slavery or practices similar to slavery, debt bondage, and forced labor. Source: U.S. Department of State archived website: <https://2009-2017.state.gov/j/tip/what/index.htm>.

Sector-geography	Thematic area that references a specific geographical area and industry and/or type of trafficking under which the Global Fund to End Modern Slavery (GFEMS) categorizes PEMS 1 projects.
Sex trafficking	When a trafficker uses force, fraud, or coercion to compel a person to engage in a commercial sex act or when a trafficker causes a child who has not attained 18 years of age to engage in a commercial sex act (Source: Statement of Work).
Stakeholder	A person or a group whose activities and interests strongly affect and are affected by the issues concerned, who have a stake in the implementation of policies, who control relevant information and resources, and whose support is needed to implement change.
Theory of change	A theory of change links inputs, activities, outputs, and outcomes to explain how and why desired changes are expected to occur. It provides clearer avenues for testing assumptions and logic underlying a strategy, logframe, or planned set of activities. (Sources: Clark, H. and AA. Anderson. 2004. <i>Theories of Change and Logic Models: Telling Them Apart</i> . American Evaluation Association 2004. Atlanta, Georgia; Vogel, I. 2012).
Trafficking in persons	When a trafficker recruits, harbors, transports, provides, or obtains a person for labor or services by using force, fraud, or coercion (Source: Statement of Work). Synonymous with modern slavery and forced labor.

EXECUTIVE SUMMARY

Evaluation Purpose and Background

The Department of State Office to Monitor and Combat Trafficking in Persons (TIP Office) launched the Program to End Modern Slavery (PEMS) in 2017 to support transformational efforts to achieve a measurable and substantial reduction in the prevalence of human trafficking (referred to as “trafficking” throughout the document). PEMS objectives, which the TIP Office developed, were designed to adhere to the spirit and letter of the Fiscal Year 2017 National Defense Authorization Act (NDAA) and relevant appropriation legislation from Congress and represented a new model to approaching trafficking for the TIP Office. The model emphasizes building evidence and showing a reduction in prevalence in key target areas through intervention-based funding of implementers. The model also includes rigorous assessments and monitoring through project-specific studies.

The TIP Office contracted with EnCompass LLC to conduct a mixed-methods process evaluation of programming and prevalence estimation studies of the Global Fund to End Modern Slavery’s (GFEMS) PEMS programming, which aimed to reduce trafficking in South and Southeast Asia. The South and Southeast Asia regions have several key characteristics that cause trafficking to thrive: large impoverished populations; major cultural, linguistic, and economic barriers for marginalized groups; porous borders; and significant populations of low-skilled labor migration, among others.

This process evaluation assessed eight GFEMS’ anti-trafficking projects still in progress, including five prevalence estimation (PE) studies spanning six sector-geographies (combination of geographical location and labor sector) across three Southern and Southeast Asian countries – India, the Philippines, and Vietnam. Sector-geographies included: internal migrant construction labor in India (1 project and 1 PE study); commercial sexual exploitation (CSE) in India (2 projects and 1 PE study); sex trafficking in Vietnam (2 projects); overseas migrant workers from Vietnam (1 project and 1 PE study); overseas migrant workers from the Philippines (2 projects and 1 PE study); and apparel industry in Vietnam (1 PE study).

Evaluation Scope, Approach and Methods

The process evaluation was launched in October 2019, with data collected between September 2020 and December 2020. It covered eight GFEMS projects and associated subawardees, and five prevalence estimation studies included in the first tranche of PEMS funding that were still underway as the evaluation began. The evaluation assessed the rigor, context appropriateness, and use by subawardees of GFEMS PE study design and implementation – a field of study in which debate remains strong among experts regarding the best methodologies to apply to various contexts to estimate often hidden, hard-to-reach trafficked populations; 2) innovatively applied an expert-informed consensus-building Delphi Technique to develop a rubric for assessing GFEMS PE study methodology design and implementation and a Prevalence Estimation Methodologies Features and Considerations (PEMFAC) tool to support decisionmaking around PE methodologies; and 3) examined project implementation and early progress in combatting trafficking by subawardees funded by GFEMS.

Throughout the evaluation, EnCompass employed participatory and appreciative approaches to design mixed-methods evaluation activities and collect and analyze data, working closely with GFEMS and the

TIP Office to ensure utility and value, while minimizing effects to interventions and beneficiaries. The evaluation team included experts in trafficking in targeted countries, human trafficking prevalence estimation, trafficking research ethics, and prevalence estimation statistics (embedded in the team and/or as consultant advisors) and engaged a 34-member Research and Prevalence Experts Panel (RPEP) to inform prevalence estimation considerations incorporated into a rubric used to assess GFEMS prevalence estimation studies and methodologies.

The evaluation faced data collection limitations. These included stronger voice from subawardees, GFEMS and its partners, and limited survivor voice. That said, sufficient data was obtained from early studies and reports to corroborate interview data and develop useful findings, conclusions, and recommendations across GFEMS prevalence estimation studies and projects implemented. We highlight a few of these below.

Learning and Recommendations from GFEMS Prevalence Estimation Studies Process Evaluation

Estimating the prevalence of human trafficking is challenging for many reasons, including the often hidden and elusive nature of the crime; stigma around the crime and its victims; legal, ethical, and definitional challenges; and growing but incomplete evidence to provide guidance about which PE methods for measuring trafficking are likely to provide accurate estimations in particular geographic and economic sector contexts.

GFEMS sought initially to fund prevalence estimation studies that would provide efficiencies and cost savings through remote data collection and building on techniques adapted from public health. TIP Office concerns about rigor led to funding several studies using more established methods. The five GFEMS prevalence estimation studies active in the evaluation period, therefore, represent a range of methods: longitudinal survey of (internal) labor migration referred to as “longitudinal migration tracking” (India construction LMT), network scale-up method and respondent-driven sampling (India Commercial Sexual Exploitation of Children (CSEC) NSUM/ RDS), longitudinal (overseas) migration tracking combined with time-location sampling (Philippines OFW LMT), and multi-stage probability proportional to size (PPS) methods (Vietnam overseas migration PPS and Vietnam apparel PPS). All PGFEMS study designs, implementation experiences, and adaptations reflected the challenges, complexities, and trade-offs associated with estimating trafficked populations or individuals vulnerable to trafficking in different contexts.

Which GFEMS methods hold promise in estimating prevalence of human trafficking across contexts or sectors?

One key evaluation subquestion sought to uncover the applicability of the methodologies to other contexts. Our assessment revealed the following:

While some GFEMS prevalence estimation study methodologies hold promise for prevalence estimation in other localities, it is not clear how generalizable these methodologies are for other sectors or contexts.

For example, the Vietnam Migration and Vietnam Apparel studies were conventional probability proportional to size (PPS) designs and key informant researchers viewed the approaches to be capable of generating valid and comparable estimates in different localities, but it may not work for all sectors (e.g., where a sampling frame cannot be developed).

The Network Scale Up Method (NSUM) + Respondent Driven Sampling (RDS) methodologies were chosen to address the specific context of Maharashtra, India CSEC, where collecting data directly from the population of interest (children) was not possible for legal and ethical reasons. This may face challenges when applied to other contexts (e.g., those that lack sufficient telephone coverage for the NSUM component or involve a poorly linked population of interest for the RDS component).

There are serious concerns about attrition and the rigor of prevalence estimates derived from the Longitudinal Migration Tracking (LMT) studies under review. The India construction study did not provide a population level estimate of prevalence, but instead provided timely information about labor circumstances for a non-representative sample of approximately 17,000 migrant construction workers. Lessons were learned about the challenges of implementing this approach from PEMS 1 studies for using this methodology to estimate prevalence in other localities, including ways to improve representativeness using previously identified population frameworks, and ways to reduce attrition from some migrant groups, including women construction workers. The evaluation team qualifies findings from PEMS 1 prevalence estimation studies based on the uniqueness of each study's COVID-19 conditions, contexts, locales, and for populations in different sectors using different methods.

Recommendations for Funders and Implementers of Prevalence Estimation Studies

Prevalence estimation study funders should ensure both experts in prevalence estimation and in human trafficking in the relevant labor industry and geograph are involved in prevalence estimation design and implementation. TIP experts should engage in discussions with researchers around challenges and decisionmaking needed to balance study breadth, rigor, contextual appropriateness, and representativeness to ensure funder objectives are met.

AND

Prevalence estimation study funders and researchers should consider the tradeoffs between requests for prevalence estimation and other learning (e.g., intervention evaluations, real-time data, and trend data) to determine the possible costs and effects that integrating these will have on prevalence estimation resources and study rigor.

Because prevalence estimation research in the TIP context is complex, requiring numerous design decisions over time, continued engagement between funders, researchers, and external experts is essential to ensuring that design and implementation decisions continue to align with the study goals and planned use of the study findings.

For example, the team concluded that all five studies experienced difficulties constructing valid sampling frames and achieving representative samples in their respective TIP contexts. Even with significant efforts on the part of research teams to do so, concessions needed to be made to balance study rigor, value, and representativeness. Challenges with sample framing (the two Vietnam studies), recruitment and attrition (India construction LMT and the Philippines overseas migrant LMT design) affected the scope or representativeness of the sample compared to the initially defined targeted population. For these and other reasons, the PEMS 1 prevalence estimation studies generally targeted specific geographic areas rather than attempting to be nationally representative. Researchers also had to make choices in operationalizing TIP definitions for their targeted populations, methodologies, and contexts. Several studies combined prevalence estimation with additional data requested to address other

research questions which sometimes posed new challenges, affected resources, and limited methods available to ensure rigorous prevalence estimation in the face of challenges encountered.

Finally, in these cases, the COVID-19 pandemic affected implementation for all five GFEMS prevalence estimation studies and will have implications for the interpretation and use of the studies, as it represents a significant change of context. In some cases, significant methodological shifts occurred due to COVID-19. In other cases, this was unnecessary due to the use of remote methods already planned and the particular population of interest, or due to the timing of data collection relative to local or national COVID-19 policies (e.g., the overseas migration study in Vietnam took place at a time when in-person data collection was possible).

Ongoing engagement with funders and experts can support researchers and funders to address challenges in ways that satisfy initial or reconsidered objectives, designs, and adaptations.

Prevalence estimation study implementers should include more thorough locally informed formative research (including gender and social inclusion assessments) and design pilot testing before undertaking a study, and the time and costs required for adaptations, to increase the likelihood of rigorous, representative estimates for the target of interest. Study funders should incorporate time and resources needed for locally informed formative research and design pilot testing into funding streams and RFPs, and researchers should include these plans explicitly in their proposals.

AND

In the short term, prevalence estimation study funders should consider the potential gains of funding the same institution to conduct multiple studies with similar objectives, methods and contexts sequentially, as learning from one study can lead to efficiencies in another. For the long term, funders should require researchers to share practical lessons learned from study implementation widely so that other researchers and institutions can benefit from this knowledge.

In addition to working with GFEMS' subawardees, in all cases, researchers collaborated with local research organizations, which, while challenging at times, provided important support in understanding context. That said, some study limitations and challenges appeared preventable had sufficient formative assessments involving survivor voice and local experts been conducted in advance of designing and implementing studies, as these could have uncovered important locally contextually important data to support more rigorous designs and practices.

For example, researchers consistently designed their studies to try to answer questions about how biological sex can affect the experiences of the populations of interest. However, researchers may have overlooked some considerations of how gender dynamics might affect study implementation, and in some cases, gender dynamics emerged that may have introduced bias to study results (India LMT, India NSUM/ RDS).

Other issues that formative research may have uncovered involved reasons later uncovered for some portions of very high early and unexplained attrition rates for the LMT study and learning of significant off-list populations needed for more accurate recruitment for the Vietnam apparel study, well after a study was under way. While some of these issues were addressed through additional studies and more complex analyses, locally informed formative research could have allowed initial designs to have incorporated this knowledge, improving efficiency of resources, reducing potential for errors associated with biases introduced, and increasing confidence in process and results.

As one kind of formative research, GFEMS prevalence estimation researchers implementing multiple studies with similar methodologies, whether more established (Vietnam Migration and Apparel using PBS) or innovative (LMT), were able to apply lessons learned from one study to the next.

This final recommendation supports all the previous recommendations.

Funders interested in more rigorous, context appropriate, and useful prevalence estimation studies—for interventions or more broadly—should fund research to address numerous questions experts in this field seek to test and explore, while balancing the need for end-user engagement and ensuring research is useful for policy action. Funders should also support carefully planned information sharing across broader, more localized groups that can benefit from and use the information to reduce modern slavery.

GFEMS researchers believe methods to estimate prevalence of trafficking need further development to improve their validity and usefulness for guiding policy and improving interventions. GFEMS' studies provide key lessons around how to operationalize humantrafficking definitions, address study recruitment and attrition, and proxy recruits. While answering several questions, GFEMS' research points to areas where further research could answer important questions about remaining key gaps in understanding. During the development of the PEMFAC, using the Delphi Technique with the 34-member expert panel, a strong level of consensus or near-consensus was achieved for numerous statements regarding the methodologies used by GFEMS and other prevalence estimation methodologies applied to human trafficking populations. During this process, additional areas for research were also identified by experts.

Funders should recognize that methodologies used in practice are not yet fully developed for human trafficking contexts to obtain valid and reliable estimates and therefore should continue to fund research to build a better evidence base on what works for prevalence estimation in these contexts. Funders should continue to support the use of diverse and innovative methods for estimating prevalence of human trafficking populations and facilitate widely sharing lessons learned from these efforts. However, funders should balance efforts to refine methodologies and answer questions about methods that are transferrable between contexts with the needs of local stakeholders in specific study contexts by engaging with local stakeholders to understand what they need from research to support policy action.

Learning and Recommendations from GFEMS Project Implementation Process Evaluation

Human trafficking is a deeply complex issue, shaped by a variety of factors that vary over geographies and sectors and may shift over time, as situations evolve. These recommendations translate findings and conclusions emerging from this assessment into concrete actions specific stakeholders can take to support more resilient programming that shows promise for affecting human trafficking.

Recommendations for Funders and Implementers of Anti-Trafficking Projects

The TIP Office and other anti-trafficking funders should fund as well as build the capacity and interest of potential partners to undertake more comprehensive, responsive anti-trafficking initiatives in challenging contexts, as these designs show resilience in implementation when faced with external challenges.

AND

Funders and implementing organizations should carefully and intentionally: 1) resource and conduct formative research and relevant assessments, including gender and social inclusion assessments, 2) engage survivors and other local actors prior to and throughout the design and implementation of anti-trafficking initiatives, and 3) where safe and appropriate, cultivate partnerships across government counterparts, civil society organizations, and private sector actors.

GFEMS' subawardees applied a multifaceted approach to combatting human trafficking, combining complementary, overlapping components with activities that engaged a diverse set of stakeholders at multiple levels and attempt to address drivers of trafficking by reducing vulnerabilities while in many cases, also increasing risks to would-be traffickers. Although evidence is limited and the COVID-19 context is unusual, this more comprehensive approach shows potential for resilience in the face of challenges and some activities employed show early potential for reducing vulnerabilities to trafficking for targeted populations and increasing risks to would-be traffickers. Projects focused on between two and four of the TIP Office's four P's – prosecution, protection, prevention, and partnerships. Those working in more of these areas, using more modes of engagement, and working with a broader array of partners could continue with some activities even when other stalled, or had to be modified or cancelled due to challenges meeting local national or US policies; challenges meeting survivor and vulnerable population's needs; or due to COVID-19 restrictions. Additionally, examples of beneficiary satisfaction appeared linked to their perceived alignment of training, supports and opportunities to specific beneficiary needs. The evaluation also yielded examples of beneficiary dissatisfaction that demonstrated a lack of this understanding.

Funders and implementers would benefit from working with and building the capacity of awardees, subawardees, and contractors to have strong locally contextualized understanding of human trafficking for the sector of interest, and to cultivate strong local partnerships. Funders can encourage these practices by requiring local partnerships in grant selection criteria, and by ensuring that funded organizations have the space and flexibility to generate relevant pre-design data, learning opportunities, and design work, including, where safe and appropriate, alongside survivors and other local partners. For funders, by further exploring modalities through which they might support organizations on the frontlines of the struggle against human trafficking/modern slavery—directly or indirectly—they could ensure their funds are going to the most impactful, best placed organizations or coalitions to address local problems and drivers of vulnerability and human trafficking. If funders were to take such actions, they might also increase the number of such organizations, improve their sustainability, and strengthen the possibility that their work will make a difference. Incorporating private partners and building their capacity to address trafficking, as PEMS has done, represents a powerful opportunity to expand this base while reducing trafficking among private for-profit businesses. Funders should support, and implementers should design and carry out key assessments and other formative research in advance of beginning projects to best understand target population needs in local contexts and, for migrants, for destination sites.

Implementing organizations working to combat human trafficking should intentionally design for and incorporate adaptive management practices into projects, including short- and long-term evaluation and learning studies, ongoing context monitoring through local networks and partners, and regular reflection and learning moments. While we have limited data that this leads to more effective programs, other studies show this to be the case, and there is evidence that implementation of GFEMS' activities benefited from these practices.

AND

The TIP Office should continue to work with implementers and other key stakeholders in ways that ensure funding and reporting timelines are adequately flexible, and implementers have the support, resources, and capacity they need to adapt to unanticipated challenges and new information, even if doing so means deviating from initial design plans.

Data from this evaluation suggest that adaptive management practices contributed to more successful implementation and positive outcomes of GFEMS subawardee projects, including by strengthening their ability to respond to unexpected challenges and to ensure approaches align to target population needs and styles. Early studies provided examples of successes across several projects (Neev Construction, India CSEC project, Blue Dragon, Fair Employment Foundation [FEF], and Blas Ople) associated with having conducted intentional studies designed to support and used for programmatic decisionmaking. Successful adaptive management, however, does not happen in isolation. Rather, as evidence from GFEMS indicates, adaptive management appears to be undertaken and used when implementing organizations are grounded in local contexts, understand the role they play in the mix of partners and toward the larger project objectives, enjoy sectoral expertise, work closely with local partners to collect and use data to inform project design and implementation, and have grant flexibility to adapt.

Implementers should optimize their adaptive capacity by investing in building their local sectoral and geographical expertise, including by working hand in hand with local networks and partners, so that they are able to identify and respond to changes in context. They should intentionally design adaptive practices, including the regular collection, reflection on, and use of data, into their ways of working, to ensure that they are able to use evidence to inform decision-making processes throughout implementation.

Funders should emphasize use of the adaptive approach reflected in GFEMS' work and TIP Office support for this work, and ensure that funding agreements and modalities provide implementing organizations with the flexibility and learning opportunities they need to identify and respond to emerging challenges throughout project implementation. This might include funding longer term projects and supporting adaptive management approaches to MEL, as well as providing for resources and training to equip subawardees and their partners with the skills they need to design and implement fit-for-purpose learning efforts. These characteristics are, to some extent, already part of, but could nonetheless be further strengthened, built out, and integrated into subsequent funding rounds.

Other Potential Uses for this Evaluation

Results of this evaluation can also inform how Congress specifies the use of funds in order to give the TIP Office more flexibility in how to manage and distribute funds. For example, the TIP Office might address timing and commissioning of prevalence estimation studies separately from anti-trafficking project implementation in order to avoid the potential complicating factors that can affect rigor and representativeness of studies. They could also fund research to help to understand the appropriate balance between prevalence estimation study designs and other information gathering that can support programs in timely ways to support adaptive management.

The report also offers considerations for future evaluations of prevalence estimation studies, for further evaluating projects and subawardees and for utilizing human trafficking definitions in prevalence estimation studies.

BACKGROUND

Trafficking in persons (TIP) is a largely hidden crime that has, over the last two decades, gained significant attention from law enforcement, human rights advocates, and policymakers. Given the heightened importance of combating human trafficking, multiple governmental organizations have begun to take action to address the problem. The U.S. Department of State's Office to Monitor and Combat Trafficking in Persons (TIP Office) has several vehicles for providing funds for special programs and projects to end modern slavery.²

The TIP Office launched the Program to End Modern Slavery (PEMS) in 2017 to support transformational efforts to achieve a measurable and substantial reduction in the prevalence of modern slavery. The first tranche of \$25 million (PEMS 1) was granted to the Global Fund to End Modern Slavery (GFEMS) in October 2017 as a 3-year grant. Through this grant, GFEMS funded projects and prevalence estimation studies in five sector-geographies: 1) India construction, 2) India commercial sexual exploitation of children (CSEC), 3) Philippines migrant workers, 4) Vietnam sex trafficking, and 5) Vietnam overseas migrant workers, as well as a research study related to the Vietnam apparel sector.

In September 2019, the TIP Office contracted EnCompass LLC to conduct a process evaluation of the GFEMS strategy for PEMS 1 subawardee and contractor efforts, including commissioned research and prevalence studies, and to identify early signs of PEMS 1 contributions to combating modern slavery. The evaluation focused on activities occurring under PEMS 1 between October 2019 and December 2020, and referenced proposed and conducted activities prior to this period where relevant (Annex 1). This report describes the design and implementation of this process evaluation, the findings that address the evaluation questions, and recommendations and considerations for further work.

PEMS 1 Mandate

Combating modern slavery is a complex challenge that requires a coordinated, multipronged strategy. The TIP Office's approach to eliminating human trafficking involves the "3Ps:" **p**rosecution, **p**rotection, and **p**revention. A fourth "P"—**p**artnerships—plays a crucial complementary role in ensuring all key segments of society are engaged in the fight to stop modern slavery. The United States applies this approach domestically and around the world. PEMS 1 objectives were designed to adhere to the spirit and letter of the Fiscal Year 2017 National Defense Authorization Act (NDAA) and relevant appropriation legislation from Congress.

In the Notice of Funding Opportunity (NOFO) developed for PEMS 1, applicants were required to create clear strategies to achieve a measurable and substantial reduction in the prevalence of modern slavery in targeted populations within partner countries (or jurisdictions thereof). Strategies were expected to complement and build on existing anti-trafficking efforts, develop sustainable local capacity of governments and civil society, and incorporate survivor perspectives. Applicants were required to present plans for robust monitoring, evaluation, and reporting and contribute data, analysis, lessons learned, and promising practices to support and amplify global efforts to combat modern slavery.

² This report uses the Palermo and TVPA definitions of human trafficking. The term "modern slavery" is also used interchangeably with "human trafficking." The exception is in discussions of prevalence estimation studies where definitions are discussed in more detail.

Key Legislative Elements of PEMS 1

The TIP Office also required the awardee to meet requirements of relevant legislation for funding of PEMS 1.³ PEMS legislation represented a new model for approaching TIP for the TIP Office. The model emphasized building evidence and showing a reduction in prevalence in key target areas, in addition to focusing on intervention-based direct funding of implementers. It also expected rigorous assessments and monitoring through project-specific studies. The legislation that allocated PEMS funding required an evaluation of the program “on a not less than biennial basis, by an independent monitoring and evaluation entity.”⁴ Underlying this approach is the idea that data-driven decision-making has the potential to create significant change and helps tell the story of how change is occurring.

PEMS 1 Awardee Expectations

The PEMS 1 awardee funding projects and contracts was expected to: 1) develop specific and detailed criteria to monitor and evaluate awarded projects; 2) implement a rigorously designed system for measuring progress against baseline data; 3) ensure that projects are evaluated on a biennial basis by an independent monitoring and evaluation entity; 4) assess prevalence in target populations and outcomes using scientifically sound, representative survey methodology; and 5) identify needs for terminating projects based on pre-determined criteria.

PEMS 1 Program

The goal of PEMS 1 is **to reduce the prevalence of modern slavery within targeted communities and industries**. The TIP Office granted PEMS 1 funding to GFEMS following an open and competitive selection process. Through PEMS 1 funding, GFEMS planned to: 1) conduct prevalence estimation studies; 2) carry out programming to reduce modern slavery, including conducting associated evaluation and learning studies developed through agendas cocreated with subawardees; and 3) leverage additional funding and expand collaboration beyond PEMS 1 efforts, to further reduce modern slavery. The first two areas are the focus of this evaluation.

PEMS 1 Prevalence Estimation Methods

PEMS 1 requires measurement of trafficking in target populations “to periodically assess progress in reducing prevalence.”⁵ GFEMS subcontracted two groups to conduct a total of five prevalence estimation studies for PEMS 1 in different sector-geographies (see Exhibit 1). GFEMS subcontracted **IST Research** to conduct four prevalence estimation studies: on the construction sector in India (LMT); overseas labor migration from the Philippines (LMT); CSEC in Maharashtra, India (NSUM and RDS); and a completed NSUM study of the apparel industry in Vietnam. GFEMS subcontracted a team from the **University of Massachusetts Lowell** to conduct two prevalence estimation studies in Vietnam: apparel industry (RDS), and forced labor among migrants to Japan and Taiwan (probability proportional to size [PPS]) (Exhibit 1, with further detail about PEMS 1 prevalence estimation studies in Annex 2).

³ Pursuant to the NDAA for Fiscal Year 2017, Sec 1298(c),

⁴ Program to End Modern Slavery (PEMS) Appropriation, FY18 Omnibus, Public Law 115–141, Sec 7060; National Defense Authorization Act for Fiscal Year 2017, Sec 1298(c).

⁵ Program to End Modern Slavery (PEMS) Appropriation, FY18 Omnibus, Public Law 115–141, Sec 7060; National Defense Authorization Act for Fiscal Year 2017, Sec 1298(c).

Strong debate exists among experts regarding the best prevalence estimation methodologies for trafficked victims overall and in different contexts. Hidden populations, including victims of trafficking, might be or feel criminalized, stigmatized, or at risk of retribution from reporting. Therefore, they can be difficult to reach when conducting research studies on prevalence. PEMS 1 studies employed several well-established methods, including multi-stage probability proportional to size (PPS), time-location sampling (TLS), and several more innovative methods for TIP populations, including respondent-driven sampling (RDS), network scale-up methods (NSUM), and longitudinal migration tracking (LMT). This evaluation assesses the rigor and context appropriateness of prevalence estimation studies funded through PEMS 1.

PEMS 1 Prevalence Estimation Methodologies

Multi-stage probability proportional to size: Selects sample elements in two or more stages, such as sampling provinces and then districts, where the probability of including each element in the sample is proportional to its known population size.

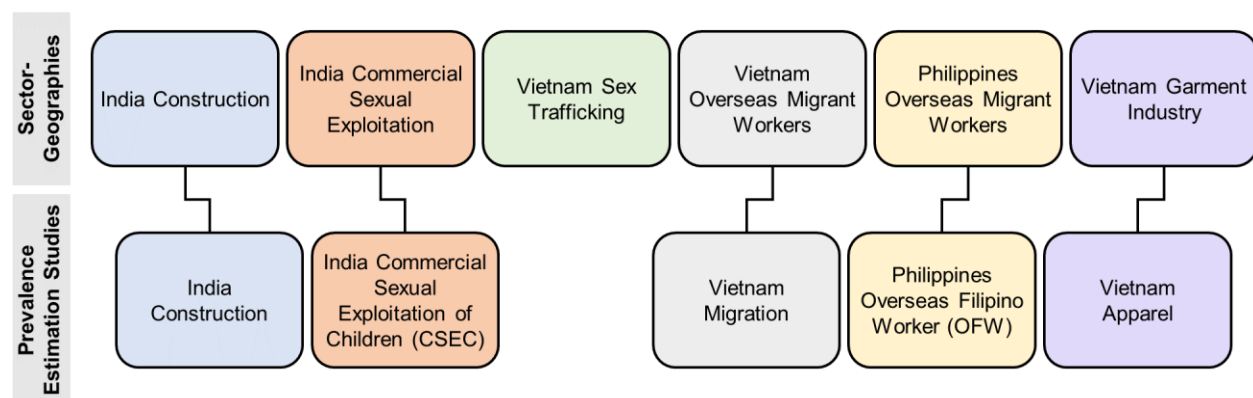
Time-location sampling: Samples individuals at target population congregation spots at discrete, predetermined times. Discrete and replicable, but time-intensive, and assumes known congregation spots are representative of population.

Network scale-up method: Surveys the general population about how many people in their network are members of the target population. Respondent biases need to be carefully modeled.

Respondent-driven sampling: Combines snowball sampling (referrals) with mathematical model weighting to compensate for the non-random sample.

Longitudinal migration tracking: Recruits a convenience or probability sample of migrants and conducts multiple follow-up surveys via telephone or SMS.

Exhibit 1: Relationships between PEMS 1-targeted sector-geographies and prevalence estimation studies



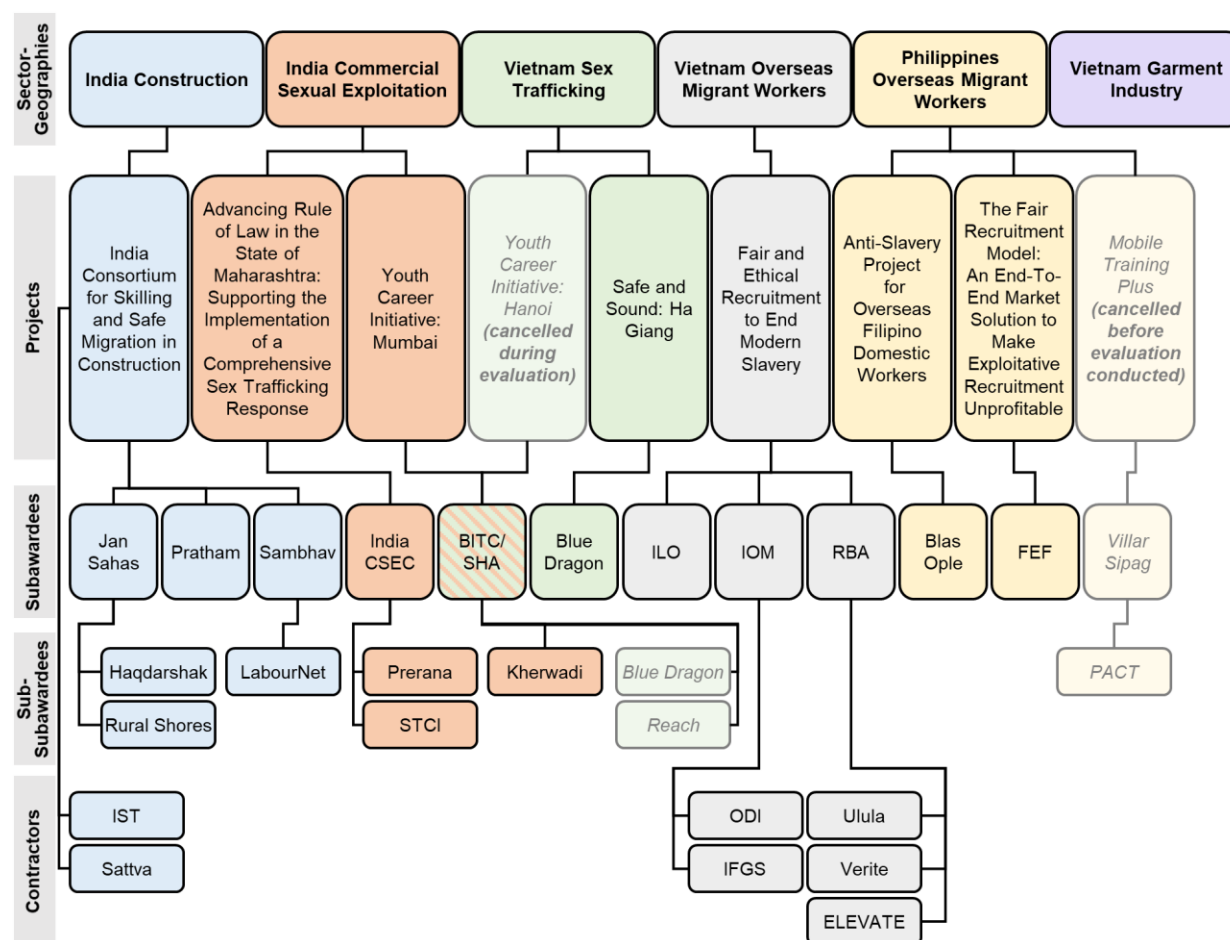
PEMS 1 Projects, Subawards, and Evaluation and Learning Studies

Projects and Subawards Associated with Key Sector-Geographies

GFEMS funded 12 subawards and one contract for work on nine GFEMS-identified projects in sex and labor trafficking. A **project** represents the level at which GFEMS has articulated theories of change and the unit of interest for GFEMS regarding impact of efforts on prevalence, and may include multiple subawards, subawardees, and contractors. Subawards and contracts will be referred to in this process evaluation report collectively as “subawards.”

Under PEMS 1, five sector-geography (all except Vietnam apparel) projects were designed with multiple components that work together to reduce the prevalence of modern slavery. Exhibit 2 is a visual representation of the relationship between sector-geographies, PEMS 1 projects, and subawardees, sub-subawardees, and contractors. As shown, one project on the overseas Filipino worker (OFW) study was subsequently canceled. This process evaluation focuses on eight projects still underway (not canceled) in January 2020 and their associated subawards, sub-subawards, and contractors, and covers 11 organizations. Further descriptions of projects can be found in Annex 3.

Exhibit 2: Relationships between PEMS 1-targeted sector-geographies, projects, and stakeholders⁶



⁶ FEF=Fair Employment Foundation, STCI=Save the Children India; ODI=Overseas Development Institute, IFGS=Institute of Family and Gender Studies

Note that three of the prevalence estimation studies are directly or tangentially connected to PEMS 1 subawardee efforts: the India construction study is integrated into India Consortium for Skill and Safe Migration in Construction (Neev Consortium) work and used by the consortium for project activities; the Philippines OFW study connects only tangentially to the work of Fair Recruitment Model (Fair Employment Foundation [FEF]) and Ople Center; and the India CSEC study covered the same sector and geography as the work of the India CSEC subawardee for the Advancing Rule of Law in the State of Maharashtra: Supporting the Implementation of a Comprehensive Sex Trafficking Response project. The prevalence studies in Vietnam examine the apparel sector and migration to Japan and Taiwan, and are not connected to the anti-sex trafficking projects being conducted through Safe and Sound: Ha Giang (Blue Dragon) and YCI Hanoi, in Vietnam.⁷

Evaluation and Learning Studies Associated with Projects

GFEMS worked with each of its subawardee project partners to design and implement an **evaluation and learning agenda** that is closely tied to the project implementation activities, and designed to validate and inform short- and long-term strategies for reducing the prevalence of modern slavery.⁸ Each subaward includes a logframe with indicators reported quarterly to GFEMS along with a narrative quarterly report, which GFEMS then consolidates into an overall PEMS 1 quarterly report submitted to the TIP Office.

Research studies corresponding to evaluation and learning questions are implemented either independently by the subawardees or in partnership with researchers, including IST Research and Nirmala Niketan College of Social Work. These studies address questions about the effectiveness of specific aspects of subawardee interventions; develop profiles of vulnerabilities to trafficking and trafficker profiles; explore how specific components of subawardee programs can be improved; obtain qualitative input from stakeholders on their experiences with the subawardee projects; or explore the sustainability, scalability, and transferability of the projects. Annex 4 provides a list of evaluation and learning studies subawardees planned or conducted.

Evidence generated from targeted research, evaluation, and subawardee monitoring, evaluation, and learning (MEL) reporting was expected to be used for learning and adaptation of project implementation by implementing **adaptive management concepts**. This process evaluation examines these adaptations, particularly adaptations resulting from the COVID-19 pandemic. In the long term, these studies are expected to contribute to modifications to PEMS 1 sector-geography theories of change and future programmatic and funding efforts. More detail on evaluation and learning studies associated with projects is available in Annex 4.

⁷ The TIP Office indicated to GFEMS it did not need to conduct a Vietnam sex trafficking study because subawardee Blue Dragon had conducted a baseline study intended to inform project design.

⁸ GFEMS PEMS 1 Monitoring, Evaluation, and Learning Strategy, Sections 1.1.1 through 1.1.4.

EVALUATION PURPOSE AND QUESTIONS

Evaluation Purpose

This process evaluation report is intended to support the TIP Office by presenting evidence of how project and subawardee efforts were implemented, challenges they experienced, strengths and facilitators that supported implementation, and to note early evidence of potential for progress for PEMS-funded awardees, subawardees, and contractors to Congress and other key stakeholders. Results from the evaluation may support resource allocation and policy decisions regarding combating international TIP. Because this is a process evaluation with a specific scope focused on subawardee work, the design does not evaluate the overall GFEMS strategy, nor specific subawardee, sub-subawardee, and contractor performance outcomes. However, the evaluation includes an assessment of prevalence estimation studies undertaken in various sectors funded under PEMS 1, and an assessment of program models or model components that may hold promise in reducing the prevalence of modern slavery in the target sectors and geographies.

EnCompass conducted the process evaluation using a mixed methods approach with two main objectives to assess:

1. The appropriateness, quality, and rigor in designing and implementing methodologies to measure modern slavery in targeted industries and geographies, as well as other research studies designed to support PEMS subawardees and contractors, and project objectives.
2. The ways PEMS 1 funding, as implemented by GFEMS, serves the target population with a focus on assessing which efforts might hold strong potential to reduce modern slavery in the targeted industries and geographies.

Reading this Report

This report describes the evaluation questions addressed, methods, findings, conclusions, recommendations, and considerations for further work. Much of the PEMS 1 program is integrated, with prevalence estimation (PE) studies linked with project implementation, and evaluation and learning studies. That said, the objectives listed above and evaluation questions (described below), cover two distinct areas: prevalence estimation and projects. The report reflects both the integrated nature of how studies and projects were undertaken and the distinct purposes for prevalence estimation studies and projects.

Methods: Data collection methods, analysis, synthesis, strengths, and limitations are summarized for PE studies and project implementation. A more detailed description of ethics and data management is included as an annex (Annex 5).

Prevalence Estimation Methodologies Features and Considerations (PEMFAC): As part of the evaluation, given the nascency and debate around methods used for TIP prevalence estimation, the team conducted an analysis and developed a tool to support decision-making around whether and how to use particular prevalence estimation methods for different types of human trafficking and in specific

contexts. This tool, along with the methodology used to develop it, is under separate cover, and helpful to review prior to reviewing the PE studies.

PE Studies: Methods are followed by a **PE Studies section** that includes reference to case summaries (Annex 6). These **case summaries** describe studies in greater detail and provide key takeaways from individual PE study assessments. Case summaries, referenced in this document and under separate cover due to length, are followed by **cross-study findings**.

Projects and Subawardees: The Projects section includes short project descriptions followed by cross-project and subawardee findings and conclusions.

Conclusions, Recommendations, and Considerations for Further Work: These sections include overall conclusions, recommendations and considerations for both the PE studies and the projects, and for funding and conducting further TIP work.

Evaluation Questions

The evaluation team developed two sets of overarching evaluation questions to address evaluation objectives. Evaluation questions were derived from the original SOW (Annex 1), analysis and discussions during the co-creation process, additional meetings with the TIP Office and GFEMS team members, and information gathered from scoping in India, the Philippines, and Vietnam in February and early March 2020.

1. How have GFEMS and PEMS 1 stakeholders designed, planned, implemented, and used prevalence estimates, interim outcome measures, and other research to improve efforts?
2. How have subawardees, sub-subawardees, contractors, and sub-contractors implemented PEMS 1 activities to combat TIP?

Each overarching evaluation question has corresponding sub-questions and lines of inquiry. The lines of inquiry provide specificity to ensure that interview questions appropriately address stakeholders' needs. All evaluation questions, including those in the original SOW, have been included as lines of inquiry.⁹

Prevalence Estimation Methodologies

Question Set 1: How have GFEMS and PEMS 1 stakeholders designed, planned, implemented, and used prevalence estimates, interim outcome measures, and other research to improve efforts?

EnCompass assessed the design and implementation of PEMS 1 prevalence estimation methodologies. For prevalence estimation studies, the evaluation's sub-questions and lines of inquiry focused on: 1) the process and rationale used to design and implement studies, including considerations and trade-offs related to scientific rigor (e.g., accuracy, precision, sensitivity), feasibility (e.g., expertise, time, cost), ethics, and adaptations to date; 2) the extent to which other PEMS 1 stakeholders were engaged in the development and implementation of studies and used study results; and 3) how GFEMS has used or plans to use these findings. Exhibit 3 outlines sub-questions and key lines of inquiry for Evaluation Question Set 1, on the PEMS 1 prevalence estimation studies.

⁹ In this report, sub-question 1.4 from the Evaluation Design has been moved to become sub-question 2.5, as it was determined to be a better fit.

Sub-Questions	Lines of Inquiry
1.1 How and why were/are prevalence estimation methodologies selected, designed, and implemented?	<p>What prevalence methodologies have been used or are planned to be used in determining baseline data and assessing potential reduction in prevalence in the future under each subaward context? When were prevalence assessments completed and when will future prevalence estimates be completed?</p> <p>For GFEMS' completed and planned prevalence studies, what was/is the process for choosing specific prevalence methodologies in each subawardee context, and why were certain methodologies chosen?</p> <p>To what extent were the processes for sample selection, developing survey instruments, and data collection scientifically rigorous and appropriate in the specific context?</p> <p>For GFEMS' completed and planned prevalence studies, what was/is the criteria and process for addressing challenges during study implementation, and what lessons were learned?</p>
1.2 How were/are subawardees and contractors involved in the design, implementation, and use of prevalence estimation methodologies and results?	<p>To what extent have subawardees and contractors had a role in the design and implementation of prevalence estimation studies?</p> <p>To what extent have subawardees and contractors used completed and/or ongoing prevalence measurement findings to inform their work?</p> <p>For incomplete prevalence estimation studies, how do the subawardees and contractors plan to use the findings?</p> <p>What other plans do GFEMS and the PEMS 1 research implementers have for disseminating and using prevalence measurement data?</p>
1.3 To what degree and in what ways has GFEMS used prevalence estimation methodologies to provide utility for PEMS 1 and future PEMS-funded efforts?	<p>What other plans for GFEMS' use of prevalence measurement data exist?</p> <p>How might the methodologies developed through GFEMS' work be applied in different localities or sectors to generate valid and comparable estimates?</p> <p>How might the prevalence assessment methods used in specific industries or communities to date be improved?</p> <p>What are lessons learned about when/where/in what context different prevalence methods can be used?</p> <p>What information is still unknown about implementation of different prevalence methods in different contexts?</p>

1) PEMS 1 Implementation, Early Progress, and Evaluation and Learning Studies

Question Set 2: How have subawardees, sub-subawardees, contractors, and sub-contractors implemented PEMS 1 activities to combat TIP?

This evaluation explored the structure and progress to date among PEMS 1 projects across India, the Philippines, and Vietnam, including potential impact in reducing modern slavery among target populations. The evaluation featured sub-questions and lines of inquiry that focused on: 1) project target populations and rationale for their selection; 2) project activities and changes over time; 3) early progress and potential for impact, including project participants' perspectives; and 4) strengths and challenges of PEMS 1 subawardee implementation to date, from both programmatic and contextual perspectives. For evaluation and learning studies, the evaluation describes evaluation and learning

studies GFEMS and subawardees have designed, implemented, and used to improve programming. Exhibit 4 outlines the sub-questions and key lines of inquiry for this second question set.

Exhibit 4: Sub-questions and lines of inquiry for Evaluation Question Set 2: Project Implementation

Sub-Questions	Lines of Inquiry
2.1 What is the target population of subawardee activities and why?	Whom are the subawardees serving: survivors of trafficking, people vulnerable to trafficking, and/or others? Why were these target populations selected?
2.2 What activities are implemented under PEMS 1, and how and why have they changed over time?	What activities are subawardees implementing under PEMS 1 to reduce trafficking? How have the activities been implemented to target survivors of trafficking, people vulnerable to trafficking, and/or others? How and why have subawardees adapted implementation of these activities over time?
2.3 What projects or implementing organizations demonstrate potential to reduce trafficking, and how are they doing this?	What project components seem to have the potential to reduce trafficking, and in what time frame? Protection/Prevention: To what extent have alumni of vocational or job skilling programs been made more secure and reduced risk of re-trafficking? Protection: From the perceptions of participants, what subawardee activities seem to have a positive impact on survivors of trafficking? On people vulnerable to trafficking? What gaps do participants perceive? Prosecution: From the perspective of key system actors within the criminal justice system, how have activities affected the criminal justice system's capacity to arrest, prosecute, and convict human traffickers? What gaps remain?
2.4 What are key strengths and challenges to PEMS 1 implementation to date?	What have been key facilitators and challenges to PEMS 1 implementation in the different local contexts (sectors, countries)? What have been key strengths and challenges of subawardee PEMS 1 implementation?
1.4 (now 2.5) What evaluation and learning questions and other research studies have been designed and implemented, and how are the results used?	What evaluation and learning questions and other research studies have been planned in each subaward context? For what purpose? When were studies completed and when will future studies be completed? For completed, in-process, and planned research studies, how have GFEMS and subawardees used results, or how do they expect to use them?

EVALUATION METHODS

Evaluation Approach and Timeline

The EnCompass approach supported the collection and analysis processes by: 1) applying multiple data collection methods, 2) employing Appreciative Inquiry techniques, 3) integrating data collection and analysis at key points in the process to support understanding of integrated aspects of PE studies and project implementation, and strongly collaborating with the awardee, contractors, and subawardees to design and achieve the sample, 4) utilizing experts to inform the design and resulting learning, and 5) applying rigorous, secure data collection, data management, and data analysis procedures.

Appreciative Mixed Methods: The mixed methods process evaluation combined purposive quantitative and qualitative data to assess prevalence estimation methodologies and subawardees' implementation efforts. The evaluation team incorporated Appreciative Inquiry strengths and vision-based approaches to data collection and analysis methods.

The evaluation team employed a mix of document review and three qualitative data collection methods to collect primary data for project implementation assessment. These included KIIs (including group KIIs), FGDs, and observations, all of which were deployed across Vietnam, the Philippines, and India, and additional individual and group KIIs were conducted by EnCompass headquarters team members. The evaluation team used a purposive sampling approach, working closely with GFEMS and subawardees to identify and engage appropriate evaluation respondents. In some cases, snowball sampling methods were also used to identify and engage additional respondents (Annex 7). The evaluation team also collected secondary data via a review of PEMS 1 and subawardee documents provided by GFEMS, subawardees, and the TIP Office. All data were collected between September 2020 and February 2021, covered activities occurring through December 2020, and were analyzed in the first half of 2021.

Collaborative and Participatory: At all stages of the process evaluation, the core evaluation team 1) worked closely with GFEMS and in-country evaluation team members to account for the complex environments in which PEMS 1 subawardees and contractors work, and 2) followed ethical guidelines and applied strong data security protocols at all stages of the evaluation process to reduce likelihood of retraumatization for survivors and to ensure the health and security of vulnerable populations, practitioners working in sensitive contexts, and during the COVID-19 pandemic.

Scoping Informed: Scoping activities—particularly an early scoping trip to India before the COVID-19 pandemic affected travel between countries—served as a useful reminder of the value of in-person visits to project and PE study activity sites to fully understand the processes subawardees and PE studies contractors and their partners use to implement PEMS 1 activities. The scoping period took place between late February and late March 2020. On-the-ground scoping activities by the international team were not possible in the Philippines and Vietnam because of the timing of the scoping period and how COVID-19 was affecting travel policies in different countries at the time. Scoping, in addition to early desk review, strongly affected the purposive sample identified during the evaluation design phase and modified through institutional review board (IRB) processes in country due to the COVID-19 pandemic.

Expert Informed: The evaluation team included a **Technical Advisory Group** comprised of individuals with expertise in prevalence estimation, TIP research, and TIP context, as part of its internal team. The Technical Advisory Group reviewed the evaluation design, was updated on progress throughout the evaluation, participated in early analysis and synthesis events and reviewed a draft of the final report.

The Technical Advisory Group initially included four individuals, but as of the data synthesis stage, there were three group members. As one returned to the TIP Office prior to the end of this evaluation, only two of the original members were available to review sections of the final report. The rubric was informed by a landscape analysis of peer-reviewed and gray literature, expert interviews, and a 34-member **Research and Prevalence Experts Panel (RPEP)**, weighing in on statements utilized for the assessment of PE studies.

Integrated: Because of the high degree of integration between several prevalence estimation studies and PEMS 1 projects, and the integrated work of IST and other contractors conducting research for project-focused evaluation and learning studies and prevalence estimation studies, there was overlap in prevalence estimation (Q1) and project implementation (Q2) primary and secondary data collection. The complete primary data collection summary can be found in Annex 7. The Landscape Analysis bibliography can be found in Annex 9. The data needed to address the particular questions and lines of inquiry were aligned through the data management and analysis steps outlined in sections below.

Prevalence Estimation Data Collection Methods and Sampling

To assess the five PEMS 1-funded prevalence estimation studies conducted in India, the Philippines, and Vietnam, the evaluation team used document review, key informant interviews (KIIs), prevalence study data review, and re-analysis, and employed the Delphi technique with a panel of prevalence estimation experts to inform the development of an assessment rubric.

Key Informant Interviews and Desk Review: The evaluation team conducted KIIs with GFEMS staff and their research partners (Annex 7) to collect primary data using a purposive sampling approach that was guided by scoping interviews with GFEMS and its partners to identify appropriate evaluation respondents. The team developed case summaries for the five PEMS 1 prevalence estimate studies, based on **document review** of GFEMS, subawardee, and subcontractor materials, as well as **statistical data analyses** of data, for the three PEMS 1 studies with sufficient data available. EnCompass first developed several interim products before developing the case summaries, including a **landscape analysis** and an assessment **rubric**.

Expert-Informed Consensus-Building around Prevalence Estimation Methodologies: The evaluation team also developed an additional product, the **Prevalence Estimation Methodologies Features and Considerations (PEMFAC)** tool, that supported the development of a rubric the evaluation team used to assess PEMS 1 studies and for future use by the TIP Office and other funders or practitioners in decision-making regarding prevalence estimation methodology selection and design. The development of the PEMFAC tool was informed by a Delphi survey process with an **RPEP** that EnCompass assembled to provide expert input. While the PEMFAC tool represents a product on its own, the process for developing the tool and the resulting statements represented a method applied for this evaluation to support the development of an expert informed rubric for assessing the use of prevalence estimation methodologies in TIP contexts.

Exhibit 5 presents a summary of methods and sample for the PEMS 1 prevalence estimates studies.

Method	Data Sources and Sample
Document Review	<ul style="list-style-type: none"> 45 documents including GFEMS prevalence estimation research contractor documents such as proposals, study designs, IRB approvals, periodic reports, and other materials related to prevalence estimation and evaluation and learning questions, and GFEMS periodic reports to the TIP Office and other MEL materials 177 published and gray literature documents on prevalence estimation methodologies for hard-to-reach and hidden populations, and completed prevalence estimations of forced labor/TIP, particularly in the sectors and geographies of interest
Online and In-Person Key Informant Interviews	<ul style="list-style-type: none"> 5 HQ GFEMS MEL and advisory staff and 1 in-country GFEMS staff 7 researchers and partners, including IST and University of Massachusetts Lowell 7 in-country prevalence estimation researchers 2 in-country learning and evaluation research partners 6 subawardee, sub-subawardee, and contractor monitoring and evaluation staff 15 selected international prevalence estimation and trafficking in persons experts
Prevalence Study Data Review and Re-analysis	<ul style="list-style-type: none"> Raw and cleaned data from the three PEMS 1 prevalence estimation studies within the evaluation period
Delphi Technique	<ul style="list-style-type: none"> 3 rounds of surveys with the 34-member RPEP

All data were collected between September 2020 and February 2021, covered activities occurring through December 2020, and were analyzed in the first half of 2021.

Project Implementation Data Collection Methods and Sampling

Document Review: The evaluation team reviewed 350 documents. These included documents shared by GFEMS, such as subawardee project proposals and narratives, quarterly reports through Q2 2020, theories of change, and MEL documentation, as well as GFEMS reporting, such as quarterly reports to the TIP Office, as detailed in Annex 7.

Key Informant Interviews and Focus Group Discussions: The evaluation team carried out a total of 162 KIIs, FGDs, and observations, with a variety of stakeholders spanning Vietnam (30 events), India (77 events), the Philippines (37), and other locations (18 events). Stakeholders interviewed included GFEMS program and project staff, TIP Office staff, subawardee program and MEL staff, sub-subawardee program and project staff, project partners, and project beneficiaries. The evaluation team, working with in-country team members, including subject matter experts, tailored each data collection instrument to suit the stakeholder groups and geographies in which they were applied. Data collection instruments were translated into local languages when necessary. With the exception of Vietnam, all KIIs and FGDs were conducted remotely because of COVID-19.

Observations: Despite the challenges posed by COVID-19, the evaluation team carried out eight observation events, as detailed in Annex 7. In the Philippines and Vietnam, the evaluation team observed several training events in person while observing all COVID-19 protocols. In India, observations of microcontractor trainings occurred remotely.

Data Analysis and Synthesis for Prevalence Estimation Studies and Project Implementation

EnCompass employed a rigorous analysis process utilizing deductive and grounded coding techniques to analyze qualitative data, systematic piloting and data quality assurance processes, and a participatory data analysis, integration, and synthesis process informed by Technical Advisory Group members.

Data Coding: The evaluation team coded data from documents, KII and FGD transcripts, and observation notes using Dedoose, an online qualitative data analysis software package. Prior to coding, the team carefully de-identified, reviewed all transcriptions/translations, and conducted necessary data cleaning and scrubbing. The EnCompass team trained all team members engaged in coding to ensure uniformity of coding and that the sources and strength of each finding were clearly indicated. EnCompass treated all responses with equal weight.¹⁰ Analysts and the codebook underwent several pilot phases to ensure coding supported development of themes aligning to sub-questions and lines of inquiry and to ensure high standards were met for intercoder reliability, validity, and accuracy.

EnCompass used a form of grounded coding, constructing initial codes according to the evaluation questions, and updating codes periodically during piloting and throughout the analysis. The evaluation team met regularly during the coding process to discuss emerging findings and reflections. Repeated ideas, concepts, and themes emerged from the review of the collected data and were fleshed out during reflection sessions.

Document Coding: For review of PEMS 1 project-related documents, EnCompass organized documents using Mendeley, which allowed for flexible organizational structures and annotation. As with the qualitative interview analysis, a system of codes was applied to categorize and summarize key information by key evaluation questions and sub-questions.

Data Analysis: After completing coding, analysts examined the data using a defined EnCompass analysis process. The process involves the examination of all excerpts and quotes coded, by evaluation sub-question and line of inquiry, to develop thematic statements and summaries of evidence, identify and include evidential quotes and excerpts, and describe data stakeholder groups and data strengths and gaps. Analysts followed the process to produce in-depth analysis summaries, by question, sub-question, and line of inquiry and key disaggregation for each project and PE study, placing these into an EnCompass analysis summary template. Human trafficking push and pull factors are different for victims and perpetrators of different ages, sexes, and genders, as are the effects of various forms of labor trafficking. Analysts also explored sex considerations, which EnCompass added to the analysis because of its importance in the TIP space. Analysts also reviewed data inductively to identify emergent themes. All analysts were trained on the process and use of internally designed analysis summary templates.

Data Quality Assurance: Data collection quality was checked in the field by the use of recordings to check against transcripts where permitted, through frequent check-ins with the evaluation team's data collection lead, through reviews of transcriptions and translations by in-country and headquarters team members, and by a carefully planned transcript cleaning and personally identifying information scrubbing process conducted prior to coding and analyzing any data.

¹⁰ Dedoose allowed for capturing and coding a wide range of qualitative data and examining them in different ways for analysis, using Microsoft Excel and Word, including the segregation of data sources to compare findings and check for intercoder reliability, validity, and accuracy.

Training, pilot testing, frequent check-ins with PE study and project implementation evaluation leads, question and response trackers and strong reviews by team leads ensured inter-analysis reliability and quality (see below). Prior to coding, the evaluation team trained all analysts on process, tools, and products. The team conducted several pilot coding trials, testing the codebook and intercoder reliability before initiating coding. The team maintained a coding question tracker to ensure issues raised during coding were shared across the team to ensure intercoder reliability as new issues arose. The team held short daily, then bi-weekly, meetings to assess progress and technical and technological issues. EnCompass' Senior Technical Advisor and/or PE study and project implementation evaluation leads conducted coding checks with 10 percent of documents, KII and FGD transcripts, and observation notes. Key team members carefully reviewed analysis summaries before synthesis events. In-country team members reviewed and provided feedback regarding their interpretations and country context, as well.

Data Integration and Synthesis: Upon completion of preliminary data analysis, the evaluation team, alongside members of the TAG, came together for a series of internal **data analysis, integration, and synthesis (DAIS) events** to share emerging themes and accompanying evidence, and to discuss, interpret, and synthesize themes into findings that cut across questions and sub-questions. Given the need for remote workshops to accomplish this work, the team utilized Zoom and Miro boards to support analytical probing and discussions of data quality and validity. The team applied an EnCompass process that involves a review of analysis summaries followed by analysts sharing themes with the team, who can probe regarding the theme meaning and evidence. An affinity mapping process is then used to identify more noteworthy themes grounded in evidence and clearly described through the data. These become the findings for the evaluation. Findings are then further discussed and affinity mapping exercises are used to identify how findings about these particular projects or studies provide more meaningful understanding of the implications of a finding or set of findings, representing a conclusion.

The team held follow-up internal sessions to develop preliminary conclusions and recommendations grounded in the data and aligned to findings. In-country team members provided subsequent input on emerging findings. The participation of in-country teams and Technical Advisory Group members helped facilitate cross-country and cross-sector comparisons, helped provide checks on data quality and strength, and ensured that data were appropriately contextualized to targeted geographies and populations.

Ethics, Data Security, Collaboration

EnCompass followed rigorous ethical standards; conducted IRB reviews in the United States and in each country where data were collected; translated informed consent forms and tools for local languages, audiences, and contexts; and managed data between countries and internally, for security and confidentiality. More details can be found in Annex 5.

Collaboration Throughout Evaluation Design and Data Collection: The EnCompass team worked closely with GFEMS, PE study contractors, and subawardees to prepare a sample in line with the initial evaluation design, and to schedule and coordinate the data collection process. GFEMS, contractors, and subawardees were very receptive and helped in providing and informing participants of interview dates. In cases where potential respondents were unresponsive, subawardees and sub-subawardees readily provided referrals to replacements. As a result, the evaluation team, working across three countries and one remote context, was, for the most part, able to achieve the sample and successfully generate high-quality data.

Sectoral and context experts were involved at every stage of the evaluation, from design to data collection to data analysis. Inputs from these experts, including members of the TAG, in-country subject matter experts, the RPEP, and members of the core EnCompass team, helped ensure that findings and recommendations accurately represented the data collected for this process evaluation, and are properly situated within and across country contexts as well as the broader anti-trafficking sector.

Adaptive and Respondent Focused: The evaluation team balanced adherence to the evaluation design, data quality, and adaptiveness to the particular context of countries and respondents throughout data collection. For example:

- When respondents were not able to scan and return signed informed consent forms via email, the evaluation team provided an option for them to submit signed forms over WhatsApp, which were then printed, scanned, and uploaded. Images were deleted at source and destination when sending through WhatsApp after being uploaded.
- The evaluation team also ensured, in all cases, adequate protection for all the documents received from participants, adhering to strictly defined data security protocols in line with the IRB requirements in the different countries.

In addition, the evaluation team trained data collectors to create and maintain an accommodating environment for evaluation respondents, even when data collection took place remotely. This meant that, in the vast majority of data collection events, respondents were accommodating and open in answering questions and sharing their opinions, insights, and suggestions. This smooth, efficient, and collegial data collection process contributed to data quality.

EnCompass included encryption and other security protocols for data collected, and factored into its timeline issues of access/interconnectivity, common time arrangement, and domestic distractions/disturbances, among others, to provide additional opportunities for interviews to occur when needed. EnCompass also conducted regular reviews and assessments of logistical plans, especially in terms of government guidelines and protocols, and the general health, safety, and well-being of researchers and participants, especially in light of COVID-19. This will be further discussed in the following section on limitations of the evaluation.

Data Analysis and Findings: Strengths of analysis and synthesis included: 1) the participatory corroboration of themes across five studies, eight projects, and multiple evaluation participants, as well as reviews by the team of feedback on findings through reviews by Technical Advisory Group members, country team members, the awardee (GFEMS) and the TIP Office; 2) coherence with external information; 3) the high level of consensus achieved through the Delphi process used to develop the PEMFAC and inform the rubric; and 4) follow-up data collection, analysis, and review by our team of statisticians conducted during analysis and review stages to fill in gaps for PE study assessments.

Evaluation Challenges and Limitations

The evaluation team faced several challenges throughout the implementation of the PEMS 1 process evaluation. These include:

The COVID-19 Pandemic

The COVID-19 pandemic was a limitation of the evaluation altogether because projects and studies stopped for periods of time during the evaluation period, and the data are incomplete to address some evaluation questions (e.g., most promising project components), as a result. More importantly, people's lives, particularly those most vulnerable to exploitation, were severely affected during the pandemic, altering 1) the existence of some projects during the evaluation period; 2) the availability of some respondents and respondent groups (e.g., government officials and construction workers); and 3) the context in which data were collected, affecting responses to interview questions.

Remote Data Collection: COVID-19 necessitated a shift to completely remote data collection in India and the Philippines, with the exception of one in-person observation event possible in the Philippines. While the evaluation team was able to undertake a successful data collection process and achieve the desired sample of stakeholders, remote data collection has its limitations. For example, connectivity issues were not infrequent, which in some cases affected the quality of KIIs. At the same time, facilitation of group KIIs and FGDs can be especially challenging in a remote context, in which interviewers are less able to see and facilitate engagement from all participants. Finally, remote data collection, in some instances, restricted participation only to respondents that could access mobile phones and/or the internet. While the evaluation team adapted the evaluation design to accommodate and reduce these limitations—such as by pivoting to audio-only interviews when video connections were not possible, switching group KIIs to individual KIIs to strengthen the comfort and safety of respondents, and by proactively adapting interview styles and methods to ensure engagement from all participants—they nonetheless limited the scope and depth of the collected data. The COVID-19 pandemic also severely limited the plans in the original evaluation design for in-person observations of project and prevalence estimation activities. In the Philippines and Vietnam, the evaluation team observed several training events in-person, while observing all COVID-19 protocols. In India, conducting observations was not possible due to COVID-19 restrictions.

Data Availability and Stakeholder Bias

Data Availability: The availability of timely and quality data is always a challenge for evaluation, particularly when a project has the size and complexity of PEMS 1. In most cases, for the document review, the evaluation team was only able to access data current through Q2 2020. Primary data collection, however, took place in Q4 2020 and Q1 2021, meaning that the evaluation team was able to combine data from the document review with more up-to-date information sourced from KIIs, FGDs, and observations during analysis. Statistical analysis was only possible for three of the five prevalence estimation studies because GFEMS partners had not collected data for two studies at the time of the request for raw and cleaned data for analysis. Additionally, critical secondary data from the awardee that was expected to replace primary data the team was requested to forego to reduce potential for retraumatization and interview fatigue of beneficiaries (e.g., beneficiary perspectives on the India CSEC subawardee's CSEC efforts) were not available during the data collection period.

Survivor Perspectives: Through scoping activities, the evaluation team identified opportunities to include the voices of trafficking survivors and vulnerable groups in data collection, where feasible and appropriate. However, scoping activities also emphasized the importance of avoiding the potential for retraumatization and fatigue (due to numerous other evaluations being conducted through projects themselves) of survivors. The necessity to minimize risks and burdens for victims outweighed the need for this first-hand perspective in some cases. Though beneficiaries from several projects participated in

data collection events (e.g., microcontractors and social service workers), few survivors were interviewed. Instead, the evaluation team collected secondary data on survivors' experiences when possible. Additionally, as noted above, planned FGDs with selected beneficiary groups (e.g., construction workers, youth participants in trainings) were not possible due to changing circumstances for these individuals and project activities associated with the timing of the evaluation relative to the COVID-19 pandemic.

Stakeholder Bias: Given that primary data collection used purposive sampling, the evaluation team expected that **selection bias** could result from the proportion of respondents who are closely tied to the project. These include but are not limited to GFEMS, subawardee staff, sub-subawardee staff, and contractor staff. Lack of critical data from subawardees during the evaluation design phase (see above) showing the relative importance of government agency perspectives to check biases of project subawardees and partners led to a design lacking sufficient data from this stakeholder group. While this is a limitation of the sampling mechanism, EnCompass attempted to mitigate this bias by conducting multiple interviews with a range of stakeholders within each project, as well as by triangulating information from different sources. That said, lack of checks from a wide range of stakeholder groups, including government agencies, other nongovernment CSOs not partnering with subawardees and working in the TIP space, and beneficiaries affected the strength of data for some projects.

The evaluation team also anticipated that **social desirability bias** would be a critical challenge for this evaluation. Social desirability bias is the tendency of survey or interview respondents to answer questions in a manner that will be viewed favorably by others. Therefore, people worried about whether their responses could jeopardize their work or overall project funding could be unwilling to tell interviewers the truth about project implementation. To mitigate this bias, the evaluation team used strict identity protections for respondents and all individuals who were involved in data collection went through rigorous training on how to collect data on sensitive issues. The training also included background of and sensitization to TIP issues in India, the Philippines, and Vietnam. While these procedures likely helped, the effects of this training were not tested.

PEMFAC

As an additional product of this evaluation, the evaluation team developed a **PEMFAC** planning tool (Annex 8). The PEMFAC tool describes methodologies for TIP prevalence estimations and best practices for use. The tool is intended to have added value and utility for future prevalence estimation planning and decision-making. The PEMFAC was developed by conducting a landscape analysis followed by implementing a modified Delphi technique^{11,12} with an RPEP that EnCompass assembled for this evaluation.

Delphi Technique

The Delphi technique is a structured communication technique that gathers data from a purposively selected panel of experts to inform a course of action, define an issue or concept, determine priorities, identify best practices, and so on. The Delphi technique is often used when informed judgement is needed on issues that are difficult to define, require significant context expertise, or are emerging fields.

Expert Panel: The evaluation team recruited for and selected a 34-member panel of international experts in prevalence estimation for TIP and other hidden populations and TIP ethics to serve on an RPEP. Short biographies of RPEP members for the Delphi technique can be found in Annex 8, Appendix 1.

Landscape Analysis: EnCompass conducted a landscape analysis that describes the use and understanding of current methodologies applied to estimating prevalence for TIP and non-TIP hidden populations. We conducted: 1) a systematic review of published and gray literature, and 2) KIIs with experts in prevalence estimation methodologies and trafficking in persons research, most of whom were RPEP members.

Preparing for First Survey Round: The team developed 355 statements on the key characteristics, strengths, challenges, and considerations for each methodology through a landscape analysis process involving document review of 177 pieces of published and gray literature, and KIIs (Annex 9). The analysis document described the use and understanding of current methodologies applied to estimating prevalence for TIP and non-TIP hidden populations. The evaluation team categorized statements by method and feature (Annex 9) and conducted three survey rounds with panel experts. Surveys were initially sent via Survey Monkey. Later rounds were sent via individual, secure Excel sheets.

Iterative Analysis: Survey responses and feedback from the panel of experts were used to find areas of consensus, identify those that lacked consensus, and hone in on ways to establish where agreement

¹¹ Fefer, Jessica P., Sandra De-Urioste Stone, John Daigle, et al. *Using the Delphi Technique to Identify Key Elements for Effective and Sustainable Visitor Use Planning Frameworks*. Published April 13, 2016. Research Article: <https://doi.org/10.1177/2158244016643141>

¹² Ziglio, E. 1995. *The Delphi method and its contribution to decision-making*. In Adler, M. and Ziglio, E. (Eds.) *Gazing into the oracle: The Delphi method and its application to social policy and public health* (pp. 3–33). Bristol, PA: Jessica Kingsley Publishers.

could be found across these areas through responses to structured surveys on key characteristics, strengths, challenges, and considerations for each methodology for the PEMFAC.

For each of the three Delphi rounds, RPEP members were sent a series of statements and asked to respond to those statements by indicating whether they strongly agree, agree, somewhat agree, somewhat disagree, disagree, strongly disagree, or have no opinion about the statement. After each survey round, statements were considered final if either of the following conditions were met:

- At least 85 percent of respondents strongly agreed or agreed with the statements
- At least 93 percent (92 percent in the third round) of the respondents strongly agreed, agreed, or somewhat agreed with the statements and at least 50 percent strongly agreed or agreed

Statements meeting these criteria in Rounds 1 and 2 were considered approved and were not included in subsequent rounds. Statements not meeting the approval criteria were revised or removed for the next round, based on RPEP feedback. After the final round, statements that did not meet the threshold for approval were either placed within the main body of the PEMFAC tool and labeled as not meeting the standard for consensus (if considered important enough to the functioning of the tool), or placed after the tool as an area lacking consensus. No statements were removed from the PEMFAC tool after the third round.

In total, over the three Delphi rounds, 478 statements were proposed, 12 were removed, 422 statements (91 percent) were finalized and included as approved statements in the final PEMFAC tool, and 44 did not meet the full standard for consensus.

Summary of Learning from the PEMFAC

The process of creating the PEMFAC showed that it was possible to achieve high levels of agreement on statements about prevalence estimation methodologies in the TIP context, although gaps in knowledge exist and some areas of disagreement remain.

The panel agreed that while definitions of TIP exist, these definitions are not precise enough to be operationalized in prevalence research.

Overall, the panel agreed that because there is no single approach that is best for all TIP contexts and while there isn't enough evidence to know conclusively which methodologies are best in different circumstances, there is enough known to guide practitioners toward methodologies that are more likely to suit particular cases. However, the panel had disagreements about when prevalence estimation is justified.¹³

¹³ Also references in PEMFAC statements GC02, GC03, GC14F.

PREVALENCE ESTIMATION STUDY ASSESSMENT

Summary of Prevalence Estimation Studies

A brief summary of GFEMS' planned PEMS 1 prevalence estimation studies and proposed methodologies for studies can be found in Exhibit 6. Planned studies included: 1) LMT study of India construction labor migrants from rural areas in Bundelkhand to Delhi; 2) LMT to assess forced labor among overseas Filipino workers; 3) NSUM and RDS methods to estimate the prevalence of minors in the sex industry in the Maharashtra region of India; 4) multi-stage PPS to measure the prevalence of forced labor among Vietnam labor migrants traveling to Japan and Taiwan; and 5) PPS to study the prevalence of forced labor in the apparel industry in Vietnam. All studies also included additional associated research questions. Case summaries (Annex 6) include more detailed descriptions of PEMS 1 prevalence estimation studies. Some proposed study methodologies changed over the course of the study in response to learning, circumstances, and challenges encountered.

Exhibit 6: Brief descriptions of PEMS 1 prevalence estimation studies and proposed methodologies

Study	Sector - Geography	Lead Contractor & Partners	Research Question(s)	Proposed Method(s)	Timeline
India Construction Workers	Indian construction workers migrating from Bundelkhand to the Delhi/National Capital Region (NCR)	IST Research Partners: Neev Consortium, Ashoka University	What is the prevalence of forced labor among internal migrants in the construction industry?	Longitudinal Migration Tracking	Data collection began in 2019, but was interrupted by the COVID-19 pandemic from March–September 2020. Completed March 2021, dissemination planned for June 2021.
Philippines OFW	Overseas Filipino workers in multiple countries, targeting countries with the most OFWs	IST Research Partners: CIFAL at the University of the Philippines, UCLA	What is the prevalence of forced labor among OFWs in Hong Kong, Japan, Taiwan, Malaysia, Singapore, Kuwait, Qatar, Saudi Arabia, and the UAE?	LMT using internet-based “river” sampling supplemented by referral-based sampling and, if feasible, face-to-face registration of migrants receiving project treatments.	Delayed by COVID-19 pandemic and redesigned. Revised design received IRB approval January 2021. Registrations began March 2021.
India CSEC	Commercial Sexual Exploitation of Children in Maharashtra, India	IST Research Partners: Kantar, UCLA	What is the prevalence of minors engaged in the commercial sex industry in Maharashtra, India?	Dual methodology: NSUM using random digit dialing and RDS with time location sampling to select RDS seeds	This study is completed. Data collection took place from March to April of 2020, and results dissemination began in October 2020.

Study	Sector - Geography	Lead Contractor & Partners	Research Question(s)	Proposed Method(s)	Timeline
Vietnam Migration	Vietnamese migrant workers from the two provinces with the highest concentrations of migrant laborers to Japan and Taiwan, Thai Binh and Ha Tinh, that had returned from working in Taiwan and Japan within the past three years	University of Massachusetts Lowell Partners: VASS, John Jay College/City University of New York	What is the prevalence of forced labor/TIP among Vietnamese migrant laborers working in Japan and Taiwan?	Multi-stage probability proportional to size sampling and qualitative in-depth interviews with selected individuals who experienced forced labor/TIP	This study is complete. Data collection began in December and was paused in March 2020 due to COVID-19. It resumed in May 2020. Results were published in October 2020, while dissemination began in March 2021 and is ongoing.
Vietnam Apparel	Workers in the garment industry in the three Vietnamese provinces with the highest concentration of garment factories, Binh Duong, Hung Yen, and Da Nang	University of Massachusetts Lowell	What is the prevalence of forced labor/TIP in the informal segment of the apparel industry in Vietnam?	Multi-stage probability proportional to size sampling and qualitative in-depth interviews with selected individuals who experienced forced labor/TIP	This was paused from March to April of 2020 due to COVID-19. Data collection began in December of 2020 and is expected to be completed in June 2021.

PEMS 1 Prevalence Estimation Studies Assessment Methodology

Several factors complicate measuring TIP prevalence: (1) trafficked individuals could be hidden during and after trafficking, so may be difficult to count; (2) many trafficked individuals are also migrants, including transnationally, which creates challenges for methods, including longitudinal studies that require tracking individuals over time; (3) documentation of trafficked individuals is often very limited, conflicting, and hard to access; and (4) methodologies being implemented have not been widely tested for trafficked populations. TIP researchers have been adapting prevalence methodologies from other sectors focused on hidden populations, including, for example: the criminal justice sector for violent criminals or traffickers of illegal wildlife or drugs, the health sector for injected-drug users or individuals with HIV/AIDS, and social sciences sector for sex workers or undocumented immigrants.

Because most PEMS 1 prevalence estimation studies were still underway during the evaluation period, EnCompass assessed the decision-making process and rationale GFEMs and its partners employed to determine how to design and implement methods, and how to address challenges in implementation to estimate prevalence of trafficking in selected sectors and geographies.

Rigor and Appropriateness Assessment Rubric

EnCompass created a rubric to assess how the prevalence estimation methodologies used in the five PEMS 1 studies were designed and implemented. EnCompass' Technical Advisory Group members and the TIP Office reviewed and provided feedback on the rubric prior to its use.

To develop the rubric, EnCompass first prepared a **landscape analysis** that describes the use and understanding of current methodologies applied to estimating prevalence for TIP and non-TIP hidden populations. EnCompass developed questions and a document management system to **systematically review** research design documents, tools, protocols, and other documentation from a range of studies. We reviewed 177 peer-reviewed and gray literature documents regarding prevalence methodologies implemented to measure hidden populations within and outside the human trafficking sphere, characteristics of these methods, and challenges faced. Documents were stored in Mendeley, a web-based platform that can be accessed via a Windows application or browser. Additional key informant interviews with 15 experts in prevalence estimation helped refine the landscape analysis and resulting rubric (Annex 10), as did survey rounds conducted with 34 prevalence experts through a Delphi technique to develop the PEMFAC (see PEMFAC, Annex 8). Annex 9 provides more details on the landscape analysis plan.

Statistical Analyses

The evaluation team's statistical group at the University of Massachusetts Statistical Consulting Center explored the treatment of missing data and outlying observations, and used relevant techniques to test estimation methods for robustness based on manipulations of the datasets themselves, as well as reviewed existing literature to compare use of these methodologies in similar and other contexts. Tests using existing data included testing weighting strategies, removing single or multiple variables or seeds, and other sensitivity analyses (see Case Summaries, Annex 6, Appendix 2 for PEMS 1 study Statistical Analysis Summaries). For the two completed studies, the consulting center attempted to replicate the researchers' results. The findings about the data and statistical methods used in each study were summarized and included in the case summaries about each prevalence estimation study.

Case Summaries

EnCompass created **case summaries** of the five PEMS 1 prevalence estimation studies. Case summaries compile information about the study design, implementation, and results. The case summaries also include the assessment regarding the rigor and appropriateness of the methodological choices of each study according to an EnCompass **rubric**, and answer other evaluation questions and lines of inquiry about GFEMS and partners' use of the study findings to date, and lessons learned from each study. EnCompass worked with a statistician and the University of Massachusetts Amherst Statistical Consulting Center to review methodologies and to attempt to replicate researchers' results and perform sensitivity tests using the research data. Case summaries are useful reading prior to reviewing findings and conclusions and can be found in Annex 6.

Case summaries include the following sections: overall summary, methodology, timeline, rationale for methodology and overall design, criteria and process for addressing challenges during study implementation, overall assessment according to the EnCompass rubric, context appropriateness as assessed by the EnCompass rubric, methodological rigor as assessed by the EnCompass rubric, ethical considerations as assessed by the EnCompass rubric, and study dissemination and future use.

Details and evidence synthesized into findings, conclusions, and recommendations are grounded in evidence outlined in the case summaries. Because of their length, case summaries have been placed in a separate annex, though they represent a key aspect of this evaluation.

Findings Across Prevalence Estimation Methodologies

This section provides findings regarding the appropriateness, quality, and rigor in designing and implementing PEMS 1 study methodologies to measure modern slavery in targeted industries and geographies. The findings address the overarching prevalence estimation studies' question: **How have GFEMS and PEMS 1 stakeholders designed, planned, implemented, and used prevalence estimates, interim outcome measures, and other research to improve efforts?**¹⁴

While case summaries provide individual study reviews and assessments, findings below represent evidence-based learning *across* studies. The analysis and synthesis processes described in the Evaluation Methods section, above, provides additional detail regarding how the team developed themes and resulting findings across studies.

Since we are discussing just five studies that were: 1) addressing different sector-geographies, 2) undertaken in different contexts, further differentiated by COVID-19 restrictions and effects on populations, and 3) conducted using different methodologies, some findings describe learning across two or more studies, and may not address all studies.

Findings attempt to show which methodologies seem feasible and rigorous, the trade-offs of rigor and cost for using studies, and suggested improvements for the future. To the extent possible, given information provided, the assessment seeks to identify which studies seem like they could be used for prevalence estimation in the future, even if methods were not successful or rigorous enough given challenges faced designing and conducting studies during this particular period with the COVID-19 pandemic.

Reading this section: Prevalence estimation study findings follow PE study evaluation sub-questions and lines of inquiry. These are identified prior to each section. PE study findings are grounded in information documented in case summaries. Readers are encouraged to read prevalence estimation study case summaries prior to reading the PE study findings in order to gain context and understanding of examples and evidence provided for the findings. The executive summary contains a list of the most relevant findings for decision-making.

Question 1.1

How and why were/are prevalence estimation methodologies selected, designed, and implemented?

The findings below address how and why the various prevalence estimation methodologies used in PEMS 1 studies were selected, designed, and implemented.

¹⁴ The original sub-question, 1.4, referring to other research conducted for projects, has been moved to question 2, as these are part of project implementation.

It includes lines of inquiry 1.1a–1.1d, which address what methods were planned, how and why methods were selected, how rigorous and context-appropriate the methods selected were, and how challenges were addressed during study implementation.

1

GFEMS sought initially to fund innovative prevalence estimation studies that would provide efficiencies and cost savings, building on techniques adapted from public health; TIP Office concerns about rigor led to funding several studies using more traditional methods. The five PEMS 1-funded prevalence estimation studies active in the evaluation period, therefore, represent a range of methods: longitudinal migration tracking, network scale-up method and respondent-driven sampling, longitudinal migration tracking combined with time-location sampling, and multi-stage probability proportional to size methods.

1.1.d. For GFEMS' completed and planned prevalence studies, what was/is the criteria and process for addressing challenges during study implementation, and what lessons were learned?

As noted in Exhibit 6 above, the methodologies employed by PEMS 1 studies include respondent-driven sampling, network scale-up methods, time-location sampling, longitudinal migration tracking, and multi-stage probability proportional to size.

How studies were selected: Studies were developed through a process involving proposed designs by researchers for particular contexts and needs articulated by GFEMS and the TIP Office, with the TIP Office requiring large structural changes to designs for at least two of the study designs.

The India Construction, Philippines OFW, and India CSEC studies were conceived as part of the first of two waves of PEMS 1 studies to provide innovative techniques adapted from public health for efficiency and cost savings through the use of remote techniques, although it should be noted that the scale and complexity of the India Construction and Philippines OFW studies meant these studies were above the average cost of prevalence studies. GFEMS funded more traditional studies in Vietnam when the TIP Office raised concerns regarding the rigor of innovative but unproven methodologies involving remote data collection.

GFEMS explained that its goal in the first wave of LMT (for migrant laborers) and NSUM and RDS for CSEC prevalence estimation studies was to be “looking for prevalence estimation techniques that are going to be new, are going to be adapted from public health, are going to potentially use some remote engagement methods, and have the possibility of doing prevalence in a more efficient way. This was confirmed in the review of a GFEMS-published blog post.¹⁵

Given this focus on innovative methods and mobile remote data collection, IST and GFEMS intended the LMTs to be pilot studies that “would be an alternative to a returnee migrant study” that enabled real-time data collection with migrant workers. IST and GFEMS understood that attrition would be a challenge and that representativeness may not be possible for the India Construction study due to the lack of a formal listing of migrant workers with which they could develop a sampling frame. The Philippines presented an opportunity to attempt an LMT with a representative sample based on official OFW demographics. The pre-COVID-19 plan was to conduct in-person TLS of migrants departing the Philippines as well as online systematic sampling through social media. Because of the pandemic, IST ultimately implemented a combination of “virtual TLS” using rosters of migrants receiving pre-departure

¹⁵ <https://delta87.org/2018/10/actionable-cost-effective-prevalence-measurement-end-modern-slavery/>

training at accredited online pre-departure orientation seminars and river sampling in the revised design. The India CSEC study similarly attempted to use innovative methods (the combination of NSUM and RDS) and via remote data collection, which was initially only planned for the NSUM component, but was ultimately used for both the NSUM and RDS surveys.

According to interviews with GFEMS, the TIP Office expressed concerns about the rigor of initial studies. For the next set of studies, GFEMS contracted with UMass Lowell to use more well-known, established approaches to conducting prevalence estimation studies. The UMass Lowell team, which GFEMS contracted to implement the Vietnam Migration and Vietnam Apparel studies, designed more traditional, probability proportional to size (PPS) surveys using probability-based sampling (PBS) that were originally intended to be nationally representative in scale. However, according to GFEMS, the TIP Office requested early in their development that the studies target specific provinces in Vietnam that were the source of large numbers of migrants to Japan and Taiwan, respectively, for the migration study and the location of large concentrations of informal and small-scale garment factories for the apparel study.

2

Each study met, mostly met, or partially met nearly all key rigor and context appropriateness criteria according to the EnCompass rubric, and each study encountered significant challenges to meeting many criteria.

Each study varied significantly in terms of targeted populations, the context in which it was designed and implemented, initial and ultimate design, implementation challenges encountered, and adaptations made in attempts to address these challenges, which limits the utility of direct comparisons between the studies regarding rigor and context appropriateness. Exhibit 7 below presents an overall assessment of rigor and context appropriate criteria for each of the five studies, based on the rubric designed for the assessment (Annex 10). Individual assessments can be found in the case summary for each study (Annex 6), which explain in more detail how each study met, mostly met, partially met, or did not meet each rubric criterion.

Exhibit 7: Assessment of the scientific rigor and context appropriateness of the five PEMS 1 prevalence estimation studies according to the EnCompass rubric

	India Construction LMT	India CSEC NSUM/RDS	Overseas Filipino Workers LMT	Vietnamese Migration PBS	Vietnam Apparel PBS
Context Appropriateness					
Appropriate literature review conducted	Not assessed, data not available	Meets	Meets	Mostly meets	Meets
Appropriate formative research and pilot testing	Partially meets	Mostly meets	Meets	Meets	Mostly meets
Defined and operationalized an appropriate definition of human trafficking	Meets	Meets	Mostly meets	Mostly meets	Mostly meets
Involved local stakeholders when appropriate	Meets	Meets	Meets	Meets	Mostly meets
Made a context appropriate decision about the use of incentives	Partially meets	Partially Meets	Mostly meets	Meets	Not assessed, data not available

	India Construction LMT	India CSEC NSUM/RDS	Overseas Filipino Workers LMT	Vietnamese Migration PBS	Vietnam Apparel PBS
Methodological Rigor					
Has well-defined target population	Partially meets	Meets	Mostly meets	Meets	Not assessed, data not available
Planned use of the study articulated before design	Partially meets	Not assessed, data not available	Not assessed, data not available	Not assessed, data not available	Not assessed, data not available
Stratification and disaggregation of data	Meets	Mostly meets	Meets	Meets	Meets
Study had appropriate plans to address relevant biases	Partially meets	Mostly meets	Partially meets	Mostly meets	Not assessed, data not available
Study had appropriate plans to satisfy key assumptions	Partially meets	Mostly meets	Meets	Meets	Not assessed, data not available
Study designed to answer questions beyond prevalence estimation*	Mostly meets	Mostly meets	Mostly meets	Meets	Meets
Plans were in place to monitor data quality	Mostly meets	Meets	Mostly meets	Partially meets	Meets
Study uses multiple concurrent methods to estimate prevalence*	Study uses single methodology	Meets	Study uses single methodology	Study uses single methodology	Study uses single methodology
Final results include confidence intervals and discussion of limitations	Not assessed, study not complete	Meets	Not assessed, study not complete	Partially meets**	Not assessed, study not complete
Ethical Considerations					
Ethical protocols minimize participant risks	Mostly meets	Meets	Mostly meets	Mostly meets	Meets
Ethical protocols minimize staff risks	Mostly meets	Meets	Mostly meets	Not assessed, data not available	Not assessed, data not available

**The TIP Office did not require PEMS 1 studies to be able to answer questions in addition to prevalence estimation or to collect data using multiple concurrent methods. Accordingly, these are considered “bonus” criteria that recognize when the PEMS 1 studies go above and beyond their mandate.*

***but can fully meet this criterion if confidence intervals and measures of uncertainty are included in a version of the study submitted for peer-reviewed publication*

3

Researchers had to make choices in operationalizing TIP definitions for their targeted populations, methodologies, and contexts.

The empirical strength of human trafficking indicators relies in part on the context in which they are used. One researcher noted that some human trafficking indicators were not as empirically strong for indicating human trafficking as one might expect, and that there are tensions between the importance of using the legal framework to determine indicators and the necessity of making human trafficking indicators sector- and context-relevant.

The approach GFEMS and its researcher partners took to defining TIP was similar in the two LMT studies. Key informants explained that staff from IST, GFEMS, and the TIP Office discussed and agreed to the indicators of forced labor, looking at both the Trafficking Victims Protection Act (TVPA) and ILO sources for definitions and indicators of human trafficking and mapped indicators to both sources that could work across populations. This was done so that:

“Down the road there wouldn’t be ... any sort of critiques on the definitions by one audience or the other, knowing that they’re a diverse field of stakeholders and people who feel strongly about definitions of forced labor and human trafficking.”

GFEMS believes that selecting indicators of forced labor that map to both the ILO and TVPA definitions will allow it to be consistent in its research going forward. As an international organization, it doesn’t want to be in a position of reporting using TVPA definitions in one study and ILO definitions in the next. By using indicators such as retention of documents, physical violence, debt bondage, wage retention, and restricted movement that map to both TVPA and ILO definitions, and tiering those indicators into categories of high-, medium-, and low-severity, GFEMS and IST sought to create a pragmatic approach that would fit the TIP Office’s needs and fits what the ILO would use as well. Contractors felt confident regarding the definitions applied to their studies.

4

PEMS 1 PE studies faced challenges with sample framing (Vietnam studies), recruitment (all studies), and attrition (LMT studies) that affected the scope or representativeness of the sample compared to the initially defined targeted population.

The two Vietnam studies were ultimately designed to determine representativeness at provincial rather than national levels. The India LMT study attempted to register a large convenience sample of 100,000 migrant workers. The goal of recruiting 100,000 registrants was in line with feedback from the Government of India that they were not interested in anything less than 100,000 registrants, and was not a statistical sample size goal. Additionally, shifts in recruitment locations and high attrition rates may have limited IST’s ability to refer to the study as representative for the targeted sector-geography. The CSEC study was able to provide a result that showed representativeness for the state of Maharashtra with several caveats involving biases introduced. Biases were introduced, for example, through gender norms with more limited phone use and responses from women in households via the random digit dialing method for data collection, the lack of male and transgender sex workers in the RDS sample, and the limitations of the triangulating RDS study.

An IST key informant noted that methods lacking a sampling frame compare to those with a sampling frame, and the challenges of having enough data to develop a sampling frame when studying trafficked populations:

“None of these are equivalent to your traditional stratified area-based sampling methods that you may be familiar with from traditional survey research. And the reason for that is that in traditional survey research, you have a frame where you know exactly how many people of different demographic backgrounds...you have all of that broken down. And with the hidden population ... maybe one of the most concise definitions of a hidden population would be that you just don’t have that ...”

There are more details about challenges each individual PEMS 1 prevalence estimation studies faced during study implementation in the case summaries for each study.

5

The COVID-19 pandemic forced PEMS 1 PE study research teams to make modifications in all five studies. The impact of COVID-19 on interpreting findings in the Vietnam and India CSEC studies is not known; the LMT studies are better positioned to address interpretation issues related to COVID-19.

The research team for both Vietnam studies expressed concern regarding the policy implications of COVID-19 for interpreting results; the relevance of findings for a post-COVID-19 era is uncertain given the widespread changes to labor migration since the studies began. Data and findings could be outdated already. As a respondent from the UMass Lowell team explained,

“We made policy implication discussions based on the data already collected which may not match the reality anymore. We don't have the confidence that the same labor practices will be continuing on as if nothing ever happened. The pandemic did happen. Our speculation is that the abuses we observed will only exacerbate in a post-pandemic stage because of the crunch and impact on employers, but this is all just speculation. We could be in the opposite direction. We don't know.”

While the Vietnam studies had very similar methodologies in practice, they differed substantially in the degree to which they were affected by COVID-19, given that the Migration study was nearing completion when the pandemic hit, while the Apparel study had only recently begun. The Vietnam Migration study was perhaps the least affected, but even the choice to use province-based data collection teams rather than a central team for this study, which was forced by the pandemic, may have contributed to an on-list/off-list variable not being recorded for part of the sample. When it became clear to data collectors that there were many off-list migrants who did not appear on the official lists they were using as a sampling frame, the data collection team decided to include them in the study but did not collect data on whether individuals were on the official lists or not for the first 22 percent of their sample.

While the India CSEC study was very different in sector and methodology, the researchers expressed similar concerns regarding their findings. The India CSEC study shifted from in-person RDS to remote RDS in response to the COVID-19 pandemic, which may have impacted the number of waves in the RDS sample. Because of the longitudinal nature of the LMT studies, COVID-19 was a particularly significant challenge, forcing a pause in data collection for the India Construction study and necessitating a redesign of the Philippines OFW sampling and recruitment strategy.

In contrast, the India Construction study was able to pivot to include a supplementary COVID-19 study using the LMT respondents, and the longitudinal nature of LMT data collection may allow the study to answer questions about the impact of COVID-19 that the other studies cannot.

6

PEMS 1 prevalence estimation researchers implementing multiple studies with similar methodologies, whether more traditional (Vietnam Migration and Apparel using PBS) or innovative (LMT), were able to apply lessons learned from one study to the next. However, the Vietnam studies and others experienced some issues that might have been uncovered by additional formative assessment phase prior to finalizing research designs and beginning pilot testing.

The UMass Lowell group was able to apply lessons learned about the challenges of on- and off-list sampling from the Overseas Migration study to the Apparel study. IST applied lessons learned from the India Construction study to the Philippines OFW study. For example, some of the strategies to address attrition experienced in the India Construction study were included in the design for the Philippines OFW study. However, the Philippines OFW study continued to experience severe problems with attrition, with a response rate of 18.6 percent in the first round, only slightly better than the 17 percent response rate for the India LMT.

In both Vietnam studies, in consultation with the Vietnam Academy of Social Sciences (VASS), the researchers originally planned to use official government lists that they believed would work well for constructing a sampling frame. During implementation of the migration study, it became clear that there were many off-list migrants who did not appear on the official lists, and the researcher team decided to include them in the study. Selecting villages by PPS according to the original sampling strategy and then modifying the strategy to include off-list migrants most likely remains a robust strategy, despite the modification. The researchers addressed the on-/off-list issues as well as they could, and including off-list respondents strengthens the study because it allows the researchers to explore whether or not there were differences between on- and off-list people. For the apparel study, UMass Lowell and its partners originally planned to use a government list as well, but the researchers realized during pilot testing that the small factories were poorly linked, making a planned link-tracing component of the study infeasible, and, as a result, incomplete, necessitating plans to find off-list factories to include in the sample, and the researchers were able to adjust plans accordingly.

GFEMS has since taken steps to ensure that appropriate formative assessment takes place prior to finalizing research designs and beginning pilot testing, and GFEMS learning from these studies influenced plans for future work:

“Now that we have a better sense of what is working and what is not working, we are able to propose cleaner designs up front for future programs, allowing for a setup of year 1 being focused exclusively on prevalence baselines that then subsequently help refine implementing partner project designs. That is what the TIP Office now requires but for a variety of reasons, that chronology was not possible with PEMS 1.”

7

For various reasons, the PEMS 1 prevalence estimation studies generally targeted specific geographic areas rather than attempting to be nationally representative.

Under PEMS 1, the India CSEC study focused on the state of Maharashtra, the Vietnam Migration study on two Vietnamese provinces, the Vietnam Apparel study on three provinces, and the India Construction study on the Bundelkhand to Delhi corridor. The two Vietnam studies were originally designed to be nationally representative, but their focus shifted to selected provinces at the TIP Office’s request. The scale of the India CSEC and India Construction studies matched the scale of PEMS 1 project implementation. The Philippines OFW study is currently designed to be nationally representative, but it remains to be seen how representative the sample will be.

“In the Philippines, it's going to be a bigger size of folks coming from a lot of different areas. So, I think we'll have to be very careful. I don't think because it will be convenience, it won't be probabilistic. At least for right now. We will limit folks making generalizations to the entirety of

the OFW population and say, ‘This is what we found based on what we have,’ but this is a convenience sample.”

There is ongoing debate in the field as to the value of targeted studies versus nationally and regionally representative studies. Certainly, both have value. One argument for more targeted studies is that the intervention programs are likely to operate at a sub-national scale, so prevalence studies that cover the same scale may be useful for informing interventions and measuring the effectiveness of programs.

QUESTION 1.2

How were/are subawardees and contractors involved in the design, implementation, and use of prevalence estimation methodologies and results?

In this section, we describe findings regarding how subawardees and contractors were (and are) involved in the design, implementation, and use of prevalence estimation data.

8

Where PEMS 1 prevalence estimation research and PEMS 1 program subawardee(s) were working in the same sector and geography, subawardees had a role in the design and implementation of the prevalence studies, which supported coordination of activities and identifying and connecting research teams to populations and networks important to designing and conducting studies.

EnCompass found evidence across both document review and interviews with key informants that researchers working on the India Construction and India CSEC studies and the Philippines OFW study worked closely with subawardees and contractors during design and implementation. For the India Construction study, subawardees played a key role in study design and implementation, and several key informants noted the critical role that Jan Sahas, in particular, played, given its on-the-ground knowledge and targeted migrant networks. For the India CSEC study, the subawardee and other local partners reviewed the initial study design and provided feedback on the NSUM survey instruments, leading to substantial redesigning to include more accessible reference groups for the NSUM instrument. For the Philippines OFW study, Blas Ople and Fair Employment Foundation were both consulted as part of the study design and play key roles in following up with flagged at-risk workers (Blas Ople) and implementing interventions (FEF).

In contrast, there were no PEMS 1 subawardees working in the same sector and geography as the Vietnam Apparel study, so no PEMS 1 subawardees had a role in the design and implementation of the study. Likewise, because there was no PEMS 1 prevalence estimation study that covered the Vietnamese sex trafficking sector-geography, Blue Dragon did not provide input into PEMS 1 prevalence estimation study design. Although the Vietnam Migration study covered the same sector and geography as the PEMS 1 work of ILO, IOM, and RBA, collaboration between UMass Lowell and these subawardees focused on dissemination, as discussed in Finding 10 below.

9

In addition to working with PEMS 1 subawardees, in all cases, PEMS 1 researchers collaborated with local research organizations, which, while challenging at times, provided important support in understanding context.

While some research teams faced challenges and, in some cases, delays associated with identifying, contracting, and collaborating with local partners, others found local partners supported moving through bureaucracies with which North America-based researchers were not as familiar. CIFAL, IST's subcontracted in-country research partner, was instrumental as the intermediary between IST and University of the Philippines for ethical approval and in reviewing and finalizing the research design, protocol, and instruments, and is the local data collection partner for the Philippines OFW study. For the Philippines OFW study, IST noted that one of the reasons for the extensive delays in getting the study off the ground was finding in-country partners that had the necessary skills and networks for the study. In India, IST attempted to work with several potential partners who ultimately determined they couldn't participate. In the Philippines, the partners' lengthy internal processes delayed IRB approvals.

Despite the challenges, all collaborating research teams found the collaboration to be highly valuable to improving designs and implementation that fit local contexts. For the India CSEC study, Kantar was involved in instrument design and carried out data collection in Maharashtra, deploying both the NSUM and RDS surveys, including developing and sampling the seeds for RDS. With India Construction, however, it was primarily the involvement of subawardees—namely, Jan Sahas, which was later accompanied by Sattva, which joined the Neev Consortium later—that assisted in important support in understanding context.

10

Most studies were not completed during the evaluation period, thereby limiting evidence of actual subawardee use of prevalence estimation study data. Study proposals and designs, reports addressing some project evaluation and learning questions, and interviews with GFEMS and project and research staff provide evidence that the India LMT studies are being used for learning and advocacy.

Two of the five PEMS 1 prevalence estimation studies were completed at the time of the PEMS 1 process evaluation: India CSEC and Vietnam Migration. One study, the India Construction study, had been ongoing for some time and collected interim data supporting project learning. The other two studies, Philippines OFW and Vietnam Apparel, were just beginning their data collection by the end of this evaluation's data collection period. Accordingly, evidence of use is limited. Early indications that interim and final results are being used for learning and advocacy include the following examples.

Although they did not provide concrete evidence, GFEMS and IST key informants believe that the India Construction study is raising awareness among migrant laborers and other key stakeholders (e.g., contractors, government labor agencies) around trafficking issues related to construction in India, and may help inform Indian labor policy groups on the development of a database for tracking migrant workers. As IST noted:

"The private sector routinely uses worker voice data to inform operational decision-making. We are confident the public and private sector will find this data useful and that it will facilitate reforms."

Key informants from IST and GFEMS believe effort should be placed into making findings from the India Construction study easier to understand for broad audiences.

There is evidence from key informant interviews with research teams (i.e., IST, VASS) as well as in-country partners that non-prevalence data supported by prevalence estimation studies has been useful, for example, through: 1) applying learning around factors associated with exploitation (Philippines OFW,

India Construction, and India CSEC studies, and Philippines OFW including the Ople Center stating in a monthly progress report that it and the IACAT secretariat are collaborating with IST to have the LMT generate statistics and information that can be used to inform policies and anti-trafficking approaches); 2) using lessons learned on contact management systems to apply to future studies; 3) lessons learned on how to design and implement longitudinal tracking prevalence estimation studies (including IST key informants noting IST's openness to sharing its data for others to use, and a prevalence research team key informant explaining that VASS in Vietnam plans to include the results of the two Vietnam studies in its database, facilitating access to other institutions for use in future studies); and 4) gaining understanding of ways pre-COVID-19 vulnerabilities affected vulnerability during the COVID-19-related lockdown in India.

With three studies left to complete, GFEMS believes it will have more examples in the coming months of how these data have been used to drive program design, advocacy, and implementation.

To address the timing issues that were apparent in PEMS 1, GFEMS also explained that the TIP Office now mandates that year 1 for all future all programs be devoted to conducting a baseline and endline study and that all prevalence estimations be mixed methods, reflecting the idea that subcontractors "will need more than a single [quantitative] number for the data to be informative for design or refinement."

QUESTION 1.3

*To what degree and in what ways has GFEMS used **prevalence estimation methodologies** to provide utility for PEMS 1 and future PEMS-funded efforts?*

This section includes findings around: 1) the degree to which GFEMS has used prevalence estimation methodologies to support PEMS 1 and future PEMS-funded efforts; 2) how methodologies could be improved; 3) how developed methodologies may be applied in different sectors and geographies to generate valid and comparable results; and 4) what information is still unknown about implementing various prevalence estimation methods. Findings includes data from PE studies assessments and the PEMFAC.

11

GFEMS has plans for widely sharing information across the field of experts, donors, and practitioners in the TIP space. Part of the PEMS 1 strategy has been to package learning and evidence study components with prevalence estimation studies. GFEMS noted the importance of "socializing" the research to a broad range of audiences.

GFEMS plans to package the prevalence estimation studies with the complementary PEMS 1 learning and evaluation research studies they supported. The U.S. State Department and GFEMS co-hosted the first of these events on February 25, 2021, with a webinar titled "Research on Commercial Sex Trafficking in Maharashtra." Nearly 60 participants joined the 11 panelists for this event. IST presented results from the PEMS 1 India CSEC prevalence estimation study, and the Dalberg Group presented results from its study on "Measuring Deterrence of CSEC in Maharashtra." The presentations offered several emerging recommendations for combating CSEC in Maharashtra. By packaging the research together, participants received a perspective of the scale of CSEC in Maharashtra, evidence about what

programmatic efforts may be effective in reducing CSEC, insights into the effects of the COVID-19 pandemic on the dynamics of CSEC in Maharashtra, and recommendations for future research and interventions.

GFEMS and the Responsible Business Alliance hosted a similar webinar with simultaneous translation for Mandarin, Vietnamese, and Japanese speakers on March 4, 2021, which featured presentations by ILO, IOM, RBA, and the UMass Lowell team. The presentations together provided a fuller picture of the drivers of Vietnamese migrant laborers' vulnerability to modern slavery. Similar to the India CSEC presentation, the prevalence estimation study set the stage for the scale of the problem and key vulnerabilities, and provided some recommendations for action. RBA elaborated on vulnerabilities and then other organizations talked about how some of this knowledge is being or will be put into action soon, including a due diligence framework and efforts to put revised Vietnamese overseas labor migration laws into effect.

GFEMS has also produced briefs summarizing the results of the prevalence studies to make the findings more accessible to a wide audience.

GFEMS prioritized disseminating the results of the prevalence studies as part of webinars, but after those are done, GFEMS also intends to support publishing them in peer reviewed journals. GFEMS also participates in METIP, the community of practice for people working in monitoring and evaluation in trafficking in persons, and shares these studies and learning from them there.

12

GFEMS plans to share learning through local actors to improve on-the-ground awareness and understanding of trafficking prevalence, and for direct use in anti-TIP interventions. It also seeks to improve awareness and understanding of modern slavery and its impacts among communities of people beyond stakeholders already focused on these issues.

A GFEMS key informant spoke of the goal of putting significant financial resources to develop evidence on prevalence issues, on intervention effectiveness issues, on systems change, and on policies issues, and to inform stakeholders and help them make evidence-informed decisions. For example, one of the ways prevalence estimation data are potentially going to be used is to inform a migrant labor policy group within India that a GFEMS staff member has been invited to join. The key informant noted that the policy group described its need for a database to track migrant workers and understand their situations and what entitlements they're getting, which, as the informant noted, "is very linked up with what we're doing." Likewise, in the Philippines, a GFEMS staff member said:

"GFEMS is in touch with government entities through other programs that are looking at how to increase touch and increase communication with OFWs once they're abroad and how to make those connections when they're in distress. So, it is definitely useful that these concepts will live beyond these studies."

Additionally, a key informant noted the importance of having a "local champion," someone who is known and respected in government circles who can champion the research work. For the India Construction study, for example, the head of Jan Sahas might fill this role.

"[What we try to do is we make it as a package, like] 'This is what we did on prevalence. And here's some other stuff we did on deterrents and rule of law and also how COVID-19 is

affecting the sex industry.’ So taken together, it builds for a really full conversation where you can talk about a lot of different aspects of the issues at once. So, we’re doing that with the Embassy. Then we also have our country coordinator who is Indian and lives in India who is doing some one-on-one meetings and also small roundtables with government entities and multilaterals like IOM, UNICEF, ILO, et cetera. One of the things that came out of a recent one, which is really helpful, is that the government entities there had asked the people of the roundtable group to be advisors on policy that they’re developing around various aspects of modern slavery, which is a fantastic in for us. And I think the way that she massages it and works some of the stuff at the local level is really helpful because she is promoting us to help as advisors to the government as opposed to ‘namers and shamers’ of an issue that is never going to get a lot of traction—positive traction.”

Further, GFEMS is working with country coordinators and partners to plan a dissemination strategy to “socialize” the India LMT research along with other learning and evaluation findings. In line with GFEMS dissemination events, one key informant noted that to date, data from the India Construction study have been technical and oriented to researchers, government officials, and implementing partners. Dissemination events will have to find ways of making the findings easier to understand for broader audiences. As a key informant from GFEMS explained, their goal is to be able to look at the India Construction study and say,

“Here are trends that are happening related to forced labor issues. And here are some of the risk factors based on that general analysis of associations and correlations.”

And further, that this study is,

“...extremely useful from a policy perspective as well. It doesn’t necessarily have to be a probabilistic sample for the government to say, ‘This is useful information that we want to do something with.’”

While the Philippines OFW and Vietnam Apparel studies are still underway, thought has been put into how the results of these studies can be used. For example, GFEMS believes the results of the Philippines OFW study could be used in pre-departure orientations to better inform OFWs about the risks they face. Additionally, GFEMS demonstrated interest in utilizing creative means to reach the public at large to improve awareness and understanding of modern slavery and its impacts, beyond stakeholders already focused on these issues.

13

There are serious concerns about attrition and the rigor of prevalence estimates derived from LMTs. The India Construction study did not provide a population level estimate of prevalence, and the ability of the Philippines Migration study to do so is unproven. Lessons can be learned about the challenges of implementing the LMT methodology from the India Construction and Philippines OFW studies practitioners could use to implement this methodology to estimate prevalence in other localities.

The rigor of LMT estimates is not yet clear. While the India Construction study did not produce a population level estimate and the ability of the Philippines Migration study to do so remains unproven, the RPEP panel’s responses to PEMFAC statements about the method indicate that there is belief among experts that the method has some potential to do so in the right circumstances. The panel agreed that “using LMT to estimate prevalence of a hidden population assumes that the research includes a

sufficiently large random representative sample at baseline and relatively little attrition” and that “probability-based longitudinal sampling is ideally best applied to contexts having large and visible migration flows.” This would seem to indicate that the contexts of the India Construction and Philippines Migrations studies are appropriate, but that the decision to use a convenience sample in the India Construction study, the challenges of attempting a representative sample in the Philippines Migration study, and large attrition in both studies are significant challenges for rigorous population-level prevalence estimates. The RPEP panel also agreed with the following PEMFAC statements about the challenges of LMTs: “appropriate and efficient sampling frames for measuring migration are usually unavailable if time-location sampling or other random sampling method is not possible.” Therefore, recruitment methods for LMT often involve a combination of non-random methods: chain-referral and targeted sampling and screening for potential candidates. These processes may violate assumptions of randomness for prevalence estimation. They also agreed that “Other studies do not attempt to have a sampling frame and follow a purposively selected sample that, unless including all of the targeted population, cannot provide prevalence estimates without significant unresolved biases.” Further, they agreed that for comparative migration studies, LMTs “usually require publicly available data and standardized questions allowing for cross-comparisons.”

GFEMS and its partners do not view the LMT methodology’s greatest strengths as generating rigorous estimates for the population of interest, especially in light of its inability to provide a population level prevalence estimate in India, or that would be comparable across localities and sectors.

“The concept of having direct engagement with workers through this worker voice scenario and getting feedback directly from them is something that the government could feasibly take up later on to say, ‘We want to have direct connection with these workers. We want to know what their situation is.’ So, I think that method is something that can be taken up from a government programming/influencing policy perspective as opposed to having to be an academic research study every time.”

GFEMS and its research partners believe the LMT approach is scalable and transferable to other contexts.

“The India LMT ... is just basically using call centers. There’s no shortage of that in India. So, I think that that’s perfectly scalable. With the Philippines one ... it’s a social media capture, which is quite simple to do. A number of organizations can do that that don’t necessarily have to be research organizations. I think they’re simple concepts that can be scaled to larger sizes.”

Another key informant noted that these specific LMT studies may not be replicable because the context will have changed so significantly.

“I cannot respond to that question because taking a pre-departure during a time of a pandemic and maybe if we try to replicate it once the pandemic is over, they’re not likely bound to give you the same set of results because of the context of the pandemic. But I think as far as design is concerned, this is generally a reasonable design. Under normal circumstances, the results should be more or less replicable. However, when you talk about things like trafficking, departure, these are subject to a lot of conditions that change over time, and given the prevailing circumstances, both at the country of origin, at the receiving country, as well as any developments in the global economy. So, there are those (factors), it cannot be replicated 100 per cent just like laboratory studies. But within reasonable bounds, I think it’s a replicable study if times were normal.”

Attrition is a key challenge in LMT designs. Key informants noted that there would be potential issues applying the entirely online approach being taken in the Philippines with populations that are not as well connected online as OFWs. Another key informant noted that it would be easier to “scale down than it is to scale up,” meaning that replicating the India Construction study in a less populous country is something that could be effective, with the caveat that it would need to be adapted to match the “cultural -isms of that country.” On the other hand, another key informant believed that the approach:

“...is scalable up. I am not enough of a sociologist/geographer to know exactly. But that was another thing about it that was interesting. That it was scalable up. Because the idea of seasonal migration for construction is a pretty widespread one and a lot of it goes on. So, in my view, it is scalable up. And the other thing is it’s an industry where you could fairly easily get an overall idea of the size of it. And so, I think if you knew you tracked X number of people and Y were number of victims, it would be easy to scale that up.”

A rationale for the LMT design best fits with labor migration. While there are doubts about the ability of the method to lead to a rigorous prevalence estimate, one of the advantages of the LMT design, as noted by key informants, is that it enables the participant to be followed up with once at their destination, particularly with remote engagement methods, and provides an option for a real-time labor migration study with the population of interest, as opposed to a returnee study or study that surveys families of migrants. As one key informant noted, “I think greatest strengths there would just be there’s really no other way to do a transnational migration study, especially if you’re dealing with multiple destination countries, unless you are remotely doing it.”

14

NSUM + RDS methodologies were chosen to address the specific context of Maharashtra India CSEC, and may face challenges when applied to other contexts.

NSUM requires that researchers have a fairly large list of known populations for the geographical area of interest for the study. Two key informants pointed out that this means that NSUM may not be possible in contexts where such statistics are not available. After initially proposing an NSUM-only design and receiving instructions from the TIP Office to make significant changes, IST and its partners selected NSUM and RDS with adult sex workers as a way of estimating CSEC in Maharashtra because those methods allowed them to estimate the population that they were not allowed to interview directly because of legal and ethical prohibitions in Maharashtra against interviewing children. For populations that are more easily accessible, it may be preferable to focus on a method that allows researchers to interview them directly rather than NSUM, even when coupled with RDS. The PEMFAC notes that:

“NSUM is not useful where human trafficking victims are not well known to the general population or are hidden from one’s acquaintances (although transmission bias can be factored into estimates).” PEMFAC statement

and

“RDS is not ideal for large, heterogeneous networks, in populations that are difficult for insiders to distinguish from the majority population, or in situations where trafficked individuals are severely intimidated and consequences for individuals found to be receiving coupons would be severe.” PEMFAC statement

and

“RDS is best suited for hidden, tightly-networked and connected populations, such as migrants speaking the same language, women or men selling sex, networked domestic workers, some forced labor, and trafficking associated with other stigmatized behavior, where individuals have access to others in similar situations.” PEMFAC statement

Key informants believed that the approach would work well in countries with similarly structured commercial sex economies, but that RDS might result in a very biased sample in a country where commercial sex is highly prevalent.

“I think you could do NSUM coupled with RDS in Kenya, Uganda, Vietnam, Indonesia, I mean the same combination would apply elsewhere. It wouldn’t work as well if the structure of the commercial sex economy was very different. It may not work in Thailand, for example, because commercial sex is very public and very prevalent so you would have to rethink how you’d do your sampling process of sex workers. RDS would maybe just give you a sort of really biased sample. Basically, RDS isn’t that great if the population is large.”

15

The Vietnam Migration and Vietnam Apparel studies were conventional PPS designs, and key informant researchers viewed the approaches taken to be applicable in different localities to generate valid and comparable estimates, but may not work for other sectors. With additional funding, or with different priorities for existing funding, respondents felt the PEMS 1 Vietnam studies could incorporate additional innovative approaches to reach more of the target population and perhaps improve reach and representativeness.

Applicability in different localities: As the UMass Lowell team noted in its research design,

“Our data collection methods follow our past best practices that are replicable and scalable. We have utilized these data collection protocols over the years and have applied and adapted in several overseas settings (e.g., India, Nepal, and Cambodia) ...”

While the methods used in these studies were viewed as conventional and transferrable by key informants, the approaches may not be easily transferred to other sectors and certain contexts. The sampling frame for the migration study relied on official lists of migrants from Vietnam to Japan and Taiwan. Such lists may be available for highly regulated migration corridors, but do not exist for many contexts, such as unofficial cross-border migration. Further, the lists of migrants proved to be incomplete in the Vietnamese context (see the Vietnam Migration case summary). Likewise, government lists of the tertiary garment factories in Vietnam were incomplete, requiring the UMass Lowell team to devise supplementary methods of sampling garment factories. The success of their approach remains to be seen, which will to a large extent determine how transferrable the approach is to other contexts.

Innovative approaches to consider: The Vietnam Migration and Apparel studies both used conventional probability proportional to size approaches to estimating prevalence. The advantage of these methods is that they are well understood. In the Vietnam Migration study, one key informant explained that although they were happy with the way the study was carried out, additional innovative approaches could be worth considering:

“I think so. I’m always up for innovating, but this one actually ended up being quite textbook, which was really nice, and offered a chance to learn some other new things that’s been put

forth out there. If I had to do something different, I might—if resources permitting and with a little bit of extra time, probably try something a little bit innovative to try to find the pockets of individuals that were suffering from forced labor exploitation through an adaptive sampling approach. So, once we find one area, we sample in and around there a little bit more to increase the overall yield of these individuals in the sample. But I will note here that adaptive sampling does pose challenges for the analyst. It's very—computationally it's intense and it's a much different estimation procedure than just typical sample mean."

Another key informant spoke to the possible value of having data collection teams gather around major holidays where it is guaranteed that many workers will be home for being contacted, for studies in industries involving internal migration.

16

Researchers supplemented quantitative information with qualitative information to provide more nuance and depth to prevalence estimation studies and suggested that further use of mixed methods studies can provide additional insights into factors affecting prevalence estimates.

While including qualitative research is generally understood as valuable in the field of prevalence research, key informants emphasized the importance of a qualitative element to any prevalence estimation study, not just as an afterthought, but as a key component from initial study design in order to understand more of the nuances of risk and protective factors, and also to ensure that the data can come alive and feel relevant and compelling when it is disseminated. Researchers noted the importance of interviews to capture changes in perceptions and knowledge around trafficking, as well as the nuances of what is happening and how people experience it. The two Vietnam studies included in-depth qualitative interviews, and the LMT studies sought to uncover nuance through longitudinal data. Although the India CSEC study did not include in-depth interviews and focused on quantitative data, GFEMS did fund largely qualitative research for a CSEC deterrence study with the Dahlberg Group and learning and evaluation studies with the India CSEC subawardee. For the India CSEC study itself, IST noted that the study could have been improved by augmenting the NSUM and RDS data with additional qualitative data, but IST and its partners chose a standard approach to NSUM that didn't model transmission bias and barrier effects to derive a minimum estimate; they argued that more innovative approaches that address these possible biases hadn't been tested or scrutinized widely. Furthermore, for the India CSEC study, one key informant explained that they scraped web forums in English and Hindi for purchasers of commercial sex (i.e., customers), which yielded a lot of information about sex worker tradecraft, where people met, and customer and sex worker strategies to adapt to the pandemic, which IST notes that in retrospect would have been valuable to include, as it may have helped interpret data during analysis.

17

Researchers consistently designed their studies to try to answer questions about how biological sex can affect the experiences of the populations of interest. However, researchers may have overlooked some considerations of how gender dynamics might affect study implementation, and in some cases, gender dynamics emerged that may have introduced bias to study results.

The PEMS 1 prevalence studies included sex (in all cases) and being transgender (in the India Construction and Philippines OFW studies) as a demographic variable and disaggregated data

accordingly to examine how they can affect the individual experiences. With the Philippines OFW study, IST planned to also include questions about gender-based violence in its survey. Ultimately, these questions were removed from the Philippines OFW study during the ethical review process at the request of local partners and the local IRB.

While IST incorporated sex as a variable when building reference groups for NSUM and when weighing responses, and the instrument recorded the sex of the participants, which allowed for disaggregated results analysis, gender dynamics played a significant role in study implementation of both the RDS and NSUM instruments, resulting in disproportionate gender representation: the RDS survey included only women, and men made up the majority of respondents for the NSUM instrument because women who answered the phone tended to pass the phone to men. Although the researchers pointed out that this is a known problem in India, EnCompass did not see evidence of a mediation plan in place to address this potential bias during data collection. In the Vietnam Migration study, researchers noted that they did not explicitly address gender in selecting and training enumerators, which one respondent thought could have affected participant responses, explaining:

“It just so happens the vast majority of our enumerators are female. If male workers overseas had experienced some gender-based abuses, would they feel comfortable revealing that to a female enumerator? It's a research methods question. I don't think anybody has actually explored that, but it's certainly a valid question. Though, it was not an issue ever reported back to us.”

Gender dynamics and the sex of individuals experiencing TIP may also have played a role in the India Construction study, one key informant noted:

“Even into the helpline if you notice women call in lesser; it is always the case, it is not because of the project or it is not because of anything, it is just that women don't call in and they don't find it comfortable to speak on the phone. Increasing socialization amongst women about the helpline, even just literacy on how to use the phone, what it means to call these are all the things we are now doing on the ground, even with female workers.”

18

PEMS 1 studies experimented with innovative methods, tools, and protocols that could support these methods to estimate prevalence in other settings.

Based on interviews and analysis of PEMS 1 prevalence estimation studies, PEMS 1 applied several innovative and remote data collection and tracking methods that appear to show potential as transferable means of estimating prevalence should key relevant criteria that were not met in the India Construction study be met. To date, these methods have had significant limitations in producing a generalizable prevalence estimate. India LMT used a call center and phone follow-ups due to low literacy rates and high cell phone penetration for the targeted population (migrant construction workers). The Philippines OFW study is using a combination of remote TLS sampling from providers of online pre-departure training with Facebook advertisements and follow-up via Facebook messenger or other social media platforms for highly connected Filipino populations. Random digit dialing (India CSEC), while not new, has transferability if working in contexts where most of the population has phones, and if phone use-related biases are addressed.

For the LMT studies, key informants believed using semi-automated (India Construction, mTracker) or automated (Philippines OFW) systems to track and follow up on people could be applied in different

localities or sectors, though it is important to ensure that the population targeted is sufficiently linked to the chosen data collection method (e.g., the India Construction study chose call center and phone follow-ups for specific reasons matching that particular population [see above]; the Philippines OFW study chose Facebook advertisements and Facebook Messenger or other social media platforms for follow-up based on knowledge around the Philippines having one of the most highly connected populations in the world, with higher literacy rates [than India's LMT target population], high rates of Facebook use, and high rates of international migration). In opting for automated data collection for the Philippines LMT, IST notes that it will assist in ensuring consistency of survey administration, use branching/filtering questions, and control for cases where participants are not eligible for the survey. IST believes that this approach will produce cleaner data than the India LMT approach, where lower levels of literacy required a call center approach. However, the initial response rate in the Philippines was very low and was only marginally higher than the India study.

Additionally, being largely tech-enabled allowed researchers to pivot easily to conduct the COVID-19 survey using the mTracker system to contact India Construction study participants. On the other hand, it is not yet clear how the disadvantages of an automated online approach that does not include human interaction (either face-to-face, by phone, or by text) may affect the quality of data in the Philippines OFW study. It is too early in this study to see if these planned innovations can overcome challenges faced in the India Construction study, although early attrition data point to similar challenges. The India Construction and Philippines OFW studies can provide lessons learned about LMT approaches, but the approach remains unproven in producing generalizable TIP prevalence estimates.

19

PEMS 1 studies illustrate trade-offs in research designs, and strengths of some studies could be considered weaknesses in others.

Timing trade-offs: One trade-off is between collecting retrospective data versus collecting real-time data. The Vietnam Migration study collects data from returned migrants, which was noted by key informants as potentially limiting the value of the results, given that laborers have already returned, whereas the two LMT studies collect ongoing data from current migrants. GFEMS believes that:

“Real-time data collection efforts such as those with the worker voice/LMT studies offer fantastic data, in real-time, that can be used by [subgrantees]. However, these types of rapid data capture with vulnerable populations are not always conducive to a more rigorous probabilistic method of measuring prevalence.”

The India CSEC study was implemented over a relatively short timeframe compared to the other studies, and the study used two complementary prevalence estimation methods to derive an estimate of CSEC in the state of Maharashtra. The study's focus is on prevalence estimation, and, accordingly, it does not gather as much data on other important questions as the two Vietnam studies, which supplement the prevalence estimation with in-depth qualitative interviews, or the LMT studies, which collect large, longitudinal datasets.

Data collection trade-offs: Another trade-off evident in the India CSEC study is that remote data collection can reduce costs and protect respondents from risks, but it may also limit the ability of a study to go into depth on additional questions beyond prevalence research because remote data collection tools need to be particularly concise to reduce the number of participants who do not complete the survey. As one key informant noted,

“Because of the hidden nature of this crime, researchers have to select the best option out of a suite of imperfect choices. In making that selection, there are trade-offs.”

Answering multiple questions: The LMT methodology has potential to answer multiple questions. One of the strengths of the LMT methodology for prevalence estimation in the TIP context is that it does not treat forced labor as static, since in reality, individuals may be in a situation that can be considered forced labor one month and not the next, and that the extent of these conditions can change seasonally. By collecting data longitudinally, LMT studies can theoretically capture the dynamism of TIP and forced labor conditions. Another potential strength of LMTs is that they may be able to assess interventions to prevent TIP in addition to estimating prevalence over time.

On the other hand, these potential strengths are mitigated by the challenges. For example, to construct a generalizable prevalence estimate, an LMT should ideally have a representative sample of the population of interest, but because IST found that “granular and recent demographic data for migrants following the specific pathway examined in this study are not readily available” (Draft Final Report), they were unable to construct a valid population-level estimate. As GFEMS notes, “more learning and collaboration amongst researchers is needed to understand how to address this issue of representativeness when a sampling frame doesn’t exist.” It is also difficult to sample migrants probabilistically. Additionally, maintaining contact with migrants over time at multiple contact points is challenging, and high nonrandom and unexplained attrition leave potential biases undescribed and unaccounted for in analysis, thereby rendering an estimate unreliable. To date, GFEMS says that it has learned that it can be challenging to reach the intended population. Also, the longitudinal designs faced significant challenges during the COVID-19 pandemic, which have been discussed earlier. Ultimately, GFEMS believes that use of remote data collection methods will only increase over time, so these pilot LMT studies in PEMS 1 can help the field understand these methods and possibly learn how to integrate them into “rigorous academic research designs.”

Trade-offs between estimating prevalence and measuring intervention effectiveness: As noted by both GFEMS and IST, the ability of the LMT study design to provide real-time data and an ongoing connection with participants makes it well-suited for testing out interventions aimed at reducing the prevalence of human trafficking and forced labor conditions.

Key informants from GFEMS and IST emphasized that an important aspect of prevalence research is to keep in mind that it is good for establishing trends over time, but not great for attributing to programs/treatments or as a program evaluation tool if you do not use randomized assignment to interventions, as there are so many other variables that are challenging to effectively take into account. In some cases, however, random assignment to interventions may not be feasible or ethical. Other key informants noted that one of the benefits of longitudinal prevalence studies is that they can measure pre/post-treatment, although they also caveated that benefit by noting that such findings are a bit nebulous as it is a non-random population and there are other variables at play. One key informant said it this way:

“I think [prevalence estimation] is more helpful in terms of giving you a very solid, ground up understanding of the problem. But in many cases, it would be misapplied to use it as kind of here's before, here's after, the treatment worked. And I think the key thing there, the key reason, being that you're not controlling for all of the confounding factors.”

Still, another key informant felt that being able to have measurements over time was still a benefit for implementing partners in prevalence estimation studies.

In both LMT studies, the idea was to integrate an evaluation of the effectiveness of interventions. For the India Construction study, the LMT was tightly intertwined with project implementation of the intervention arms, and each adaptation of an intervention has potential repercussions on the LMT study. Maintaining clear documentation of adaptations and revisions helps account for confounding variables during analysis. Although IST provided some documentation on the adaptations to the LMT studies, information was missing; however, IST provided useful and detailed information on adaptations for the India CSEC study. For the Philippines OFW study, due to significant delays in beginning data collection for the study, the timeline for intervention implementation and the LMT study are not aligned.

To test the effectiveness of interventions, an LMT would ideally have random assignment to intervention and control arms, although this was never part of the India Construction study LMT design, which did not randomly assign participants to receive interventions and seeks to determine correlation between receipt of interventions and reductions in conditions of forced labor, not causation. However, GFEMS cautioned against assuming that random assignment to intervention arms is required for understanding the effectiveness of interventions, noting that there “are a number of different approaches one can take to understanding effectiveness and many modern slavery interventions are not well-suited for a traditional RCT with random assignment.”

Nevertheless, with the adaptations and unexpected challenges faced during the design, implementation, and early analysis of the LMT studies (the last two only for the India LMT), the shorter-term goal of being able to use an LMT to provide direct interventions to vulnerable populations should be weighed against the potential drawbacks of biases that may be introduced by including such interventions. For example, in the India Migration study, the need to ensure that participants enrolled in interventions necessitated changing the source of registration from the source of migration in Bundelkhand to the destination in Delhi NCR. For prevalence estimation, it would have been preferable to enroll migrants at source because by registering them at destination, it is possible that migrants who suffer forms of trafficking immediately upon arrival may be missed.

20

The PEMS 1 researchers believe methods to estimate prevalence of trafficking in persons need further development to improve their validity and usefulness for guiding policy and improving interventions.

Studies could be improved through further developments, including additional testing of methods and recent improvements made to methods within these contexts (India CSEC), and understanding more about migrants on and off official lists (Vietnam Migration).

In the India CSEC study, IST assumes that biases are both systematic and static, affecting both baseline and endline estimates the same and making trend analysis feasible. IST calls for more research into testing and evaluating which methodological improvements to the study would be reliable and effective. IST made conservative estimates of prevalence. When any results are uncertain, IST stated that it believes conservative estimates have greater value than overestimates. IST points out that the ability to control for factors such as transmission bias, barrier effects, and recall bias when performing NSUM estimation is still limited, and some new attempts to adjust for these factors may risk over-parameterization, yielding unpredictable and volatile predictions as a result of overfitting to known population data. For this reason, further testing and evaluation are necessary to validate whether (and which) methodological improvements attempted in the study are reliable and effective.

In the case of the India CSEC study, a KII respondent noted their skepticism about the rigor of RDS, stating that,

“I am not entirely convinced that respondent-driven sampling is as rigorous as the literature says that it is. It is very widely used. It's used in public health. I think it's been brought over from public health into these hidden population studies. I'm not convinced that it's really as rigorous of a sampling process as professors and researchers who have written about it think. ... I think we didn't answer that, but it wasn't really our charge to answer that.”

While the Vietnam Migration study used a conventional probability proportional to size sampling design, there were several aspects of the study that raised questions: 1) the proportion of people solicited for interviews and those refusing to participate, and any effects this has on representativeness; along with this, any nonrandom biases (e.g., age, gender, source location) associated with refusal to participate; 2) in some cases, “off-list” respondents accounted for 33 percent of respondents in a village, but this varied from place to place, which may have had implications for the proportionality-to-size sampling calculations.

Conclusions for Prevalence Estimation Studies

CONCLUSION 1. PEMS 1 study designs and adaptations reflected the challenges, complexities, and trade-offs associated with estimating populations of trafficked persons and individuals vulnerable to trafficking in different contexts, and when studies are integrated with interventions.

Aligned to Findings 1, 2, 3, 4, 7

The five PEMS 1 study designs reflect choices that are specific to each study context. Even the studies that might initially appear relatively similar had significant differences. The Vietnam Migration and Vietnam Apparel studies both use probability proportional to size sampling using government lists to construct a sampling frame and use a similar approach to operationalizing a definition of human trafficking, but also differ in important ways. For example, the Migration study is retrospective, asking migrants about their experiences overseas within the last three years, while the Apparel study asks about current conditions, which requires different procedures to address ethical considerations; the government lists in the Migration study contain individual migrants to be sampled, whereas the lists in the Apparel study contained factories, and were less complete. The India Construction and Philippines OFW studies share features including how they operationalize a definition of human trafficking, but differ considerably in their scope (internal migration versus transnational migration), sample sizes, and method of conducting follow-up surveys, to name a few. The India CSEC study used a unique NSUM+RDS approach that was tailored to the India CSEC context, where interviewing minors directly was not possible. The methods chosen for each study fit for the circumstances of each study: PPS where government lists held promise for a valid sampling frame in Vietnam, LMT where tracking migrants and to integrate research into programmatic activities, and NSUM with the general population and RDS with an adjacent population in the India CSEC context, where the population of interest could not be interviewed directly.

Tailoring methodological approaches to the study context has significant advantages compared to attempting to use a single methodology in all contexts. It allows studies such as the two LMT studies to be integrated with project implementation providing near real-time data to support decision-making to affect migrant labor experiences, potentially reducing trafficking.

Conclusion 2: Constructing a valid sampling frame and achieving a representative sample is difficult in TIP contexts. Even with efforts on the part of research teams to achieve these goals, concessions need to be made to balance study rigor, value, and representativeness.

Aligned to Findings in 2, 3, 4, 6, 7, 13

PEMS 1 studies were initially designed to pilot innovative remote methods for a TIP context and advance the knowledge of the TIP field while also producing lessons learned. Remote techniques offer cost effective means for reaching mobile populations such as domestic and international migrants, on the one hand, but there is little evidence in social science research that these methods produce valid results. Attrition rates for the remotely-run LMT studies were high and led to questions about the representativeness of the sample. The Vietnam studies were designed to use more established methods, but off-list migrants and factories created challenges for their sampling frames, reflecting the challenge of constructing valid sampling frames for hidden and hard-to-reach populations. These challenges were largely mitigated in the Vietnam Migration study and may prove to be in the Vietnam Apparel study as well. Overall, reaching and creating a representative sample of the intended hidden population was a challenge in all cases, as reliable sampling frames don't exist for hidden populations such as domestic migrants in India's construction sector. That said, there is research using a variety of prevalence methods in the TIP context and in other fields from which lessons can be learned. As GFEMS notes, "more learning and collaboration amongst researchers is needed to understand how to address this issue of representativeness when a sampling frame doesn't exist."

Conclusion 3: Studies demonstrated the importance of funding and resourcing formative assessments and pilot testing, including gender assessments, for understanding study context and successfully implementing prevalence estimation research. Studies also benefited from involving local partners in the study design and implementation.

Aligned to Findings 2, 6, 8, 9

Studies showed that taking the time to conduct formative assessment and pilot testing of tools is important to the successful implementation of prevalence estimation research in the TIP context. Through pilot testing, UMass Lowell and its partners were able to determine that planned link tracing sampling was not feasible and that many factories were missing from the government list they planned to use to create their sampling frame, and were able to adjust accordingly. On the other hand, pilot testing and formative research did not uncover the off-list issues for the Migration study, leading to challenges they had to address in the data collection and analysis phases. Formative assessments could also have helped the India CSEC study formulate a plan to address potential gender bias in phone call responses and refine their RDS incentives plan, which may have improved on the RDS results in which most responses came from the initial seeds and first round of referrals, and only produced responses from nine of 36 districts in Maharashtra. Likewise, formative research and pilot testing for LMT could have uncovered issues with survey length more quickly and reduced high rates of early attrition, and might have supported all studies in operationalizing TIP definitions. All five PEMS 1 prevalence estimation studies involved U.S.-based lead researchers working closely with local partners. Local involvement supported connecting researchers to critical networks and identifying contextual issues that required consideration. Although not evaluated, GFEMS also intends to work through local actors to improve on-the-ground awareness and understanding of TIP prevalence within communities, and for direct use in anti-TIP interventions.

Conclusion 4: The COVID-19 pandemic affected implementation of the five PEMS 1 prevalence estimation studies and will have implications for the interpretation and use of the studies, as it represents a significant change of context.

Aligned to Findings 4, 5, 10, 11, 12, 19

The COVID-19 pandemic forced changes in the implementation of studies that were underway (India Construction, Vietnam studies) and modifications or the complete redesign of studies that had not yet started (Philippines OFW). Researchers and their partners worked to minimize the effects of these challenges in cases like the Vietnam Migration and India CSEC studies where COVID-19 led to delays in data collection and required modification of data collection plans, and worked to understand the effects of COVID-19 in the India Construction context.

In all cases, COVID-19 has implications for interpreting the findings of studies, even in the cases where study implementation was less affected by the pandemic, because it is reasonable to assume that the COVID-19 pandemic will lead to significant changes in the study contexts: CSEC in Maharashtra may shift to be more online, and patterns of labor migration from Vietnam to Japan and Taiwan may change as a result of shifting economic and logistical factors caused by the pandemic.

While these factors may mean that the estimates these studies produce may quickly be out of date, GFEMS believes the completed PEMS 1 prevalence studies and the complementary learning and evaluation studies GFEMS and its partners disseminate provide useful insights into how the pandemic has affected these dynamic human trafficking contexts.

Conclusion 5. While some PEMS 1 prevalence estimation study methodologies hold promise for prevalence estimation in other localities, it is not clear how generalizable these methodologies are for other sectors or contexts. Even where contextual cross-sector transferability was clear (LMT), the method was less favorable as a prevalence estimation method than as a means for ongoing and timely engagement with workers.

Aligned to Findings 6, 13, 14, 15, 16, 19, 20

As noted in Finding 14, the methodology applied to the India CSEC study made sense for the context and could work well in countries and regions with similarly structured commercial sex economies. NSUM requires large lists for known populations for a given geographical area of study (e.g., for adult sex workers in a given area), and RDS is best suited for hidden, tightly-networked and connected populations. Thus, NSUM may not be possible in some contexts, such as cities or regions where such statistics are not available, and RDS might not apply to networks cut off from one another, such as for some migrant agricultural laborers and domestic workers.

The sampling methodology applied to the Vietnam migration study, involving official lists, was found to be problematic and it is unclear how successful alternative approaches introduced to overcome gaps in lists, by sampling garment factories directly, will be; this will have a bearing on transferability. The methodology would not apply in areas where such lists are not available or are even less complete.

Remote data collection using NSUM and LMT represented a cost-effective method for conducting larger studies with hard-to-reach populations (India Construction, India CSEC). A number of global shifts in use and power of technology increase the applicability and scaling potential these methods have, even with some of the challenges in bias potentially introduced or faced. The prevalence estimation methodologies themselves show that the data collection method may overcome some barriers, but

cannot supplant the support of sampling frames if available, and robust triangulated designs, in the absence of a sampling frame.

The rigor of prevalence estimates derived from LMT studies remain unclear. Representativeness in the India Construction study is challenged primarily by attrition and non-representative sampling. The Philippines OFW study has attempted to mitigate these issues, but is currently underway and the success of these efforts remains to be seen. While the prevalence estimates derived from LMTs are likely subject to significant limitations, the scale of the India Construction study LMT may lead to greater acceptance of results. The ability to track changes over time and provide timely assistance to workers in forced labor conditions are benefits of the methodology.

Conclusion 6. All studies faced important limitations and challenges, but these studies provide key lessons around how to operationalize TIP definitions, address study recruitment and attrition, and proxy recruits.

Aligned to Findings 3, 13, 14, 15, 16, 17, 18,

All of the PEMS 1 studies have important limitations, challenges, and lessons learned. For example, IST was transparent about the challenges it faced with attrition and recruitment for the India Construction study, and challenges with obtaining local IRB approval for an LMT design in the Philippines, although additional documentation of these challenges and adaptations could provide valuable lessons learned. Some examples of lessons learned include the need for automating the process of ensuring that at-risk participant follow-up is conducted, rather than sending information to partners on a bi-weekly basis. Revisions to tracking systems represent important details that could be helpful to future implementers of LMT designs. PEMS 1 researchers have been transparent about the challenges and limitations of their studies and lessons learned.

Operationalizing TIP definitions: GFEMS believes that selecting indicators of forced labor that map to both the ILO and TVPA definitions will allow it to be consistent in its research going forward. As an international organization, it faces challenges if in a position of reporting using TVPA definitions in one study and ILO definitions in the next. By using indicators such as retention of documents, physical violence, debt bondage, wage retention, and restricted movement that map to both TVPA and ILO definitions, and tiering those indicators into categories of high-, medium-, and low-severity, GFEMS applied this approach to both the studies it conducted with UMass Lowell and those it conducted with IST as partners.

While this overall approach provided clarity and cross-context compatibility of results, the researchers still needed to make specific choices of how to operationalize indicators of forced labor. For example, in previous studies using this approach, UMass Lowell had asked questions about ability of workers to leave their place of employment without fear. In the specific context of the Vietnam Overseas Migration study, workers' visas are tied to their place of employment, so the questions UMass Lowell had used in previous studies to assess this indicator needed to be changed.

These specific challenges reflect the overall challenge of balancing accuracy in a specific context with comparability across different contexts.¹⁶

¹⁶ This concept was also confirmed in the PEMFAC: GC102) Different contexts make standardized measures difficult, both for legal and research purposes. Nevertheless, the use of standardized measures and datasets could improve comparability of research output.

Applying learning: That PEMS 1 researchers implementing multiple similar studies were able to apply lessons learned from one study to the next reflects the challenges of implementing prevalence estimation research with hidden and hard-to-reach populations and the value the lessons learned could have for the field. While PEMS 1 India CSEC researchers transparently described how to improve study design and implementation, PEMS 1 LMT researchers were not able to meet a key assumption of longitudinal studies—that of maintaining contact with LMT participants—and have yet to be able to explain high rates of nonrandom attrition while well into the study, thereby weakening confidence in results. In the future, LMT studies could better identify strategies for probabilistic sampling of migrants to allow for more generalizable statements.

The PEMFAC tool created as a product of this evaluation outlines numerous best practices for the implementation of the prevalence estimation methodologies employed in the PEMS 1 studies and other methodologies. There was consensus among the PEMFAC panelists about most of these best practices.

Conclusion 7: While answering a number of questions, PEMS 1 research points to areas where further research could answer important questions about remaining key gaps in understanding, including ways to operationalize definitions of TIP, addressing challenges of remote recruitment and data collection, applying hybrid methodologies, managing and accounting for attrition, and validating sampling frames.

Aligned to Findings 3, 15, 19, 20

The PEMS 1 researchers implemented studies using more traditional PPS designs and innovative remote LMT, NSUM, and RDS approaches. While showing that these methodologies can be implemented in these contexts, questions remain about the rigor of prevalence estimates derived from LMT studies, the transferability of NSUM+RDS to other contexts, and the feasibility of constructing reliable sampling frames for PPS in TIP contexts. Some of these researchers believed that more innovative approaches should be incorporated into these kinds of study designs to test novel methodologies and find effective ways to reach the hardest to reach parts of hidden and hard-to-reach populations. Overall, the methods to estimate prevalence of trafficking in persons need further testing and development to improve their validity and usefulness for guiding policy and improving interventions. The PEMFAC process brought to light a number of areas where further research would be useful for developing and refining prevalence estimation methodologies for the TIP context that reflect the experience of the PEMS 1 studies. Areas needing further research include:

- All studies: Operational definitions of TIP¹⁷
- LMT and other methodologies: Remote data collection best practices¹⁸

¹⁷ This area of research aligns with the following PEMFAC statements:

GC90) The extent to which different operational definitions of measures lead to higher or lower estimates of prevalence requires further research. The development of more precise, consensus definitions may be able to alleviate this issue.

GC94A) Work is needed to identify cross-country differences in human trafficking terms and definitions with the goal of delivering policy recommendations that unify and unambiguously specify definitions so that consistent terms are used for studies, policy, law enforcement and other relevant purposes.

¹⁸ This area of further research aligns with the following PEMFAC statement:

PBS44) Further research is needed to determine effective practices in undertaking surveys with smart phones and social media in various local contexts and types of trafficking. In what circumstances do smart phones and internet-based surveys facilitate access to hard-to-reach populations and in which do they exclude those without access?

- Hybrid methodologies¹⁹
- Representative samples for LMT²⁰
- Attrition for LMT²¹
- Validated sampling frames for PBS²²

¹⁹ This area for further research aligns with the following PEMFAC statement:

PBS45A) Further research is required on developing hybrid methodologies where probability-based strategies can be used in combination with other, innovative strategies.

²⁰ This area for further research aligns with the following PEMFAC statement:

LMT59) More probability-based LMT that follows labor migrants to their destinations is needed to support improved understanding of design and sampling needs, attrition mitigation, and appropriate non-random attrition-related weighting schemes.

²¹ This area for further research aligns with the following PEMFAC statement:

LMT60) Learn what incentives work to minimize nonrandom attrition, depending upon goals of the longitudinal study (length/ frequency/type of data collection) and target respondent characteristics/ needs.

²² This area for further research aligns with the following PEMFAC statement:

PBS17A) A sampling frame is difficult to construct for almost any well-hidden population. An accurate sampling frame is required for probability-based methods to be implemented. For this reason, probability-based methods should only be used to estimate hidden populations in combination with innovative strategies. This could include qualitative evidence that could inform a solid sampling frame.

FINDINGS ACROSS PROJECT IMPLEMENTATION ASSESSMENT

Summary of Projects

As noted previously, GFEMS funded nine projects that aimed to address sex trafficking and labor trafficking in India, the Philippines, and Vietnam. Each project provided different forms of support to populations vulnerable to trafficking, and/or survivors of trafficking, including domestic workers, overseas labor returnees, and sex trafficking victims. Exhibit 8 briefly summarizes each subawardee project and its sector-geography, highlights key populations served, and explains project activities. One project, focusing on Philippines migrant labor, was canceled. This process evaluation therefore focuses on the eight projects still underway as of January 2020, covering 11 organizations across five sector-geographies. More detailed descriptions of projects are available in Annex 3. Throughout the findings, we refer to projects by the subawardee name.

Exhibit 8: PEMS 1 subawardee projects

Project	Subawardee	Sector-Geography	Description
India Consortium for Skilling and Safe Migration in Construction (India)	Neev Consortium: Jan Sahas Sambhav Pratham Education Foundation	India Construction (Internal Migrant Labor)	Identifies and registers migrant construction workers, tracking them through their migration journey. When TIP victims are identified through this system, Neev works with local government to rescue survivors. Works to ensure delivery of entitlements to workers. Trains and builds capacity for microcontractors. Certifies worker construction skills through Recognition of Prior Learning.
Youth Career Initiative (YCI)/ Empowering Survivors and Building Youth Resilience against Trafficking through Hospitality Sector Skills Development (India)	BITC/SHA	India Sex Trafficking	Works in Mumbai to establish a scalable and replicable model to support youth survivors of trafficking to sustain freedom and at-risk youth to prevent trafficking by brokering relationships between NGOs and hotels/ hospitality businesses. Sustainable Hospitality Alliance—which took over the project in 2020—now provides training and sensitization to hotels/ hospitality, while supporting NGOs with guidance and resources.
Advancing Rule of Law in the State of Maharashtra: Supporting the Implementation of a Comprehensive Sex Trafficking Response (India)	India CSEC subawardee	India Sex Trafficking	Works to rescue and restore CSEC victims by supporting rescue operations, providing survivor services, and strengthening systems of rehabilitative care. Strives to increase prosecution by working with police agencies, judiciary, and local government agencies to provide training on trafficking and support prosecutors through the trial process.
Safe and Sound Ha Giang (Vietnam)	Blue Dragon	Vietnam Sex Trafficking	Works to improve rule of law and increase arrests and convictions of traffickers by providing training to legal aides and lawyers. Aims to sustain freedom for sex trafficking survivors by supplying access to livelihoods training, employment services, trauma care, and social worker support. Also, strives to promote prevention and build resilient communities by raising awareness through work with schools and community-based organizations.

Project	Subawardee	Sector-Geography	Description
Youth Career Initiative (YCI) (Vietnam)	BITC Vietnam	Vietnam Sex Trafficking	Works in Hanoi to establish a scalable and replicable model to support survivors of sex trafficking to sustain freedom and at-risk youth to prevent trafficking by brokering relationships between NGOs and hotels. BITC provides training and sensitization to hotels while supporting NGOs with guidance and resources.
Fair and Ethical Recruitment to Combat Modern Slavery (Vietnam)	ILO Consortium: ILO IOM RBA	Vietnam Overseas Migrant Labor	Seeks to conduct research, pilot activities, and mapping to inform future investment in improving recruitment agencies' practices and providing more transparent labor supply chain mapping to ethical employers.
The Fair Recruitment Model: An End-to-End Market Solution to Make Exploitative Recruitment Unprofitable (Philippines)	FEF	Philippines Overseas Migrant Labor	Works to build an end-to-end ethical recruitment solution for migrant workers by placing and sourcing workers using an ethical model, providing pre-migration training for workers, and promoting ethical recruitment as a viable business model.
Anti-Slavery Project for Overseas Filipino Domestic Workers (Philippines)	Blas Ople	Philippines Overseas Migrant Labor	Promotes ethical recruitment by establishing a new, dedicated, multi-sector task force to address TIP cases that involve forced labor or trafficking of OFWs. Creates an integrated case management system (ICMS) to coordinate information, involve stakeholder agencies, improve detection, investigation, and case-building, and increase prosecution. Promotes ethical recruitment practices based on and verified using the ICMS. Also aims to establish reintegration programs for victims.

This section presents findings that assess the ways PEMS 1 funding, as implemented by GFEMS, serves the target population, with a focus on determining which efforts, based on the collected evidence, might hold strong potential to reduce modern slavery in the targeted industries and geographies. The overarching question to which the findings below correspond is: **How have subawardees, sub-subawardees, and sub-contractors implemented PEMS 1 activities to combat TIP?**

As described in the Limitations section of this report, it should be noted that the evaluation team had access to only limited data from project beneficiaries and government counterparts. The presented findings therefore reflect the collected evidence, but may not always comprehensively capture the views of all stakeholder groups affected by subawardee projects.

The presented findings are organized by evaluation sub-question and address relevant lines of inquiry. We have taken the liberty of rearranging the original sequence of sub-questions to help show, through emergent findings, how the story of PEMS 1 project implementation unfolds.

2.1 What is the target population of subawardee activities and why?

GFEMS identified that domestic workers, overseas labor returnees, and sex trafficking victims constitute a large proportion of the global population of TIP victims. GFEMS therefore selected subawardees that developed activities targeting these groups.

1

GFEMS funded PEMS 1 subawardees in India, Vietnam, and the Philippines working with domestic workers and overseas labor returnees who were vulnerable to trafficking and/or trafficking survivors, and sex trafficking victims. GFEMS conducted formative research showing that these populations account for a large percentage of TIP victims globally.

GFEMS had a particular interest in funding organizations working in conditions that might be favorable for making a measurable difference with respect to trafficking within selected sector-geographies and target populations. GFEMS conducted research, used relevant criteria, and developed a process for accepting competitively submitted bids for work efforts.

Process for identifying projects: GFEMS conducted research on 25 countries, mapping and scoring each against the selected criteria in a country-industry dashboard. This review revealed that sex trafficking victims are the largest population of TIP victims globally, closely followed by domestic workers. Overseas labor returnees in the domestic work, construction, and manufacturing sectors accounted for over half of private sector TIP victims globally. GFEMS therefore decided to focus PEMS 1 awards on organizations that worked with these target populations, within sector-geographies in which local conditions—including the political environment, government support, and the capacity of NGOs and civil society working on human trafficking issues—were seen as potentially favorable for making a measurable difference against human trafficking.

Criteria used to identify projects: Evidence from internal documents provided by GFEMS, interviews with GFEMS and subawardee staff, and discussions with GFEMS indicate that, early in PEMS 1, GFEMS and the TIP Office jointly agreed on 11 criteria for use in selecting geographies and sectors for intervention under PEMS 1. These criteria were:

1. Scale and nature of modern slavery
2. Donor agreement
3. Ease of PEMS ramping up
4. Current political sensitivities
5. Ability to complement GFEMS' current country efforts
6. Transnational approach needed
7. Opportunity to leverage corporate partners on the demand side
8. Existing industry initiatives
9. Magnitude of industry
10. Proof of concept
11. Susceptibility to trade incentives, regional competition

2.2 What activities are implemented under PEMS 1, and how and why have they changed over time?

PEMS 1 subawardees implemented a variety of activities that sought to combat trafficking in persons, including by reducing the vulnerability of target populations, strengthening reintegration and support services, and building the capacity and willingness of criminal justice system actors to investigate and prosecute trafficking cases. Many subawardees adapted activities over time in response to various challenges and opportunities, but largely adhered to initial theories of change. In describing activities,

the evaluation includes proposed ways GFEMS and subawardees believed activities would help survivors and/or reduce trafficking or re-trafficking of vulnerable populations. In this section, these represent their beliefs and not the evaluation team's assessment of this potential.

2

Most subawardee projects included longer-term, complementary activities involving NGOs, CSOs, and other non-state actors, and intended to reduce vulnerabilities, strengthen anti-trafficking support services, and/or support support reintegration.

Long-term, complementary sets of activities: According to project theories of change and project reports, all projects (except for ILO Consortium) included a range of complementary activities. In combination, subawardees intended these activities to address human trafficking by reducing the multiple, intersecting vulnerabilities faced by target populations, enhancing the quality and scope of anti-trafficking support services available to those populations, and supporting reintegration.

For example, Neev Consortium skills validation and certification through Recognition of Prior Learning (RPL) for migrant construction workers and engagement with businesses to connect microcontractors with new contractors and beneficiaries were intended to strengthen employment opportunities and reduce vulnerabilities; site visits, through which the project aimed to ensure adequate working conditions for intended beneficiaries, provided additional anti-trafficking support. Reports show BITC/SHA in India and YCI – Hanoi in Vietnam (BITC) took the same approach of building a network of engagement and support among the nonprofits, hotel employers, and community groups in an attempt to generate employment opportunities while providing reintegration services to training participants.

Reports from Blue Dragon indicate that the project also worked through multiple pathways. Blue Dragon sought to strengthen anti-trafficking support services by training criminal justice system actors on how to identify, investigate, and prosecute human trafficking cases, while also providing emergency assistance and support to reduce vulnerability and support reintegration. Blas Ople likewise strengthened anti-trafficking support services by implementing an integrated case management system (ICMS), while supporting reintegration by coordinating agency and NGO resources to provide services directly to returned OFWs and survivors. FEF and the India CSEC subawardee also implemented complementary sets of activities.

Working with locally based CSOs: Subawardees appear to have coordinated closely with NGOs and civil society actors throughout the implementation of PEMS 1 to broaden and deepen survivor access to anti-trafficking services, including legal support, identification of victims, case management, social and after-care, and alternative business, work, and livelihoods support. In some cases, coordination took the form of trainings provided by subawardees, through which subawardees sought to build the capacity of NGOs and CSOs to provide services to trafficking survivors and those vulnerable to trafficking. In other cases, subawardees developed guidance and resources on anti-trafficking policies and interventions with NGOs.

For example, quarterly reports and interviews with subawardee staff indicate that the India CSEC project sensitized and trained NGO service providers, helping to them to better understand how to connect youth survivors to a variety of support services, including legal aid, income generation support, and counseling programs. Similarly, with a caveat that this is based on reports and interviews from Blue Dragon's own staff, the project trained social workers in the provision of trauma-informed psychological care, so that social workers could do a better job of working with and assisting trafficking survivors in Vietnam. Through the network of trained social workers, Blue Dragon reports that the project has

helped survivors access healthcare, housing, livelihoods assistance, and vocational training. Blue Dragon has early indications that its long-term approach to training and mentoring appears useful:

"I think a lot different organizations are now focusing on trauma-informed care and trying to improve the skills of social workers. But... [Blue Dragon's trainings recur throughout] a year and with the mentoring approach in-between and so walking along side with the social workers as they do their work. And you know we haven't been able to do a full evaluation of that yet but it seems to have been quite successful."

Project reports indicate that Blas Ople developed a variety of resources for non-governmental stakeholders to use, including a multi-sector Task Force Against Trafficking of OFWs, as well as the ICMS and accompanying guidance. CSOs used these resources to identify and track exploited workers, initiate rehabilitation and reintegration assessments, and refer survivors for after-care services. The theory of change, proposal documents, and quarterly reports of BITC/SHA suggest that this project also helped NGOs engage with partners in other sectors, brokering partnerships between shelters and anti-trafficking organizations to protect survivors, and private sector businesses such as hotels, and developing guidance and resources to assist NGOs in maintaining productive working relationships with hotels. The Neev Consortium, BITC Vietnam, and ILO Consortium also coordinated with NGO and CSO actors in their efforts to support survivors and those vulnerable to trafficking.

3

Several projects focused efforts on mapping drivers of trafficking in different contexts (Neev, ILO, India CSEC, and Blue Dragon), and directly raising the risks faced by human traffickers (India CSEC, Blue Dragon, and Blas Ople), especially training and sensitizing prosecutors and police to human trafficking, and increasing awareness within communities.

Mapping drivers: Half of the PEMS 1 projects implemented activities designed to explore and map the human trafficking landscape, including drivers of trafficking in different contexts (Neev, ILO, India CSEC, and Blue Dragon). GFEMS reports note their approach: to better prevent human trafficking, anti-trafficking advocates must first understand the factors that drive the vulnerability of those at risk of trafficking, as well as the conditions that encourage human traffickers.

Proposals and quarterly reports reveal that the Neev Consortium conducted LMT surveys with migrant construction workers to collect evidence on the prevalence of trafficking and trafficking risks at a sectoral level. The ILO Consortium used mapping to determine the viability of ethical recruitment initiatives in Vietnam, identify various stakeholders that affect recruitment, and assess their engagement and buy-in for ethical recruitment. The India CSEC project, according to quarterly reports, carried out a buyers arrest project, which involved conducting research that examined the possibility of buyer arrests as a strategy to increase legal and social consequences for trafficking, as well as a deterrence study that aimed to measure and analyze the behavior, perceptions, and decisions of CSE buyers and traffickers in Maharashtra. Proposal documents and quarterly reports indicate that Blue Dragon included several research activities aimed at understanding the trafficking landscape, including an expert review of national referral mechanism policies, vulnerability surveys and assessments, and survivor vulnerability and trafficker profile reports. The timing of study completion and report publication fell outside of the evaluation period, for the most part, and this report does not assess the quality of studies or resulting reports.

Increasing risks to traffickers: According to proposals, theories of change, and quarterly reports, the India CSEC subawardee, Blue Dragon, and Blas Ople, in line with the TIP Office’s 3 Ps framework, developed and implemented activities that sought to make it riskier for human traffickers to engage in trafficking. The India CSEC subawardee, for example, sought to bolster the rule of law in various ways, including by supporting ways to increase risk to potential traffickers:

- Investigations of buyers of sex with minors
- Raid and rescue operations
- Training for police and prosecutors
- Equipping and training child welfare committees
- Working directly with community members

Documents from Blue Dragon suggest that the project sought to raise the risks for traffickers by training legal aid staff and lawyers, providing legal representation, and collaborating with police to rescue and repatriate adult and child victims of trafficking. Blas Ople supported the implementation of an ICMS, through which anti-trafficking stakeholders in government and civil society were helped to track the progress of potential trafficking cases, thereby raising the likelihood that cases of human trafficking will be prosecuted. A GFEMS interviewee, speaking about the ICMS, noted:

“... in the short term, it's going to vastly improve the way that they deal with trafficking victims on a systems level and really help the coordination between the different agencies...especially in prosecution and the protection and reintegration spheres.”

Addressing the criminal justice system: Two projects explicitly sought to equip criminal justice system actors—lawyers, prosecutors, police, and others—with the skills and capacity to better support at-risk individuals and trafficking survivors. Based on proposals, theories of change, and narrative reports, the India CSEC subawardee worked with legal aid clinics to provide legal services to trafficking survivors, and trained police officers, prosecutors, and others to identify, investigate, and prosecute trafficking cases more effectively. A respondent from the India CSEC subawardee reported that training was effective. With regards to police training, a respondent noted:

“... they are interested, they are keen, they in-fact are pursuing us to continue provide training, so then they are asking for the curriculum, even now we have got a letter from the AGD training, he has asked us to talk to the Maharashtra Police Academy...the police are really, really interested and whatever of our experiences we have seen with the police, wherever we have trained, they have gone two or four or five steps ahead of us and actually have utilized that training in the cases. ... So, we are certainly seeing police, HDA officers and certain local police stations taking a very active interest in you know pursuing investigations and trying to bring down the networks.”

A deterrence study conducted by GFEMS included qualitative interviews and quantitative surveys of 1,678 respondents, as well as analysis of web and online data to uncover and identify supply and demand side trends for use of technology in CSEC using observable online sources.²³ The study showed that potential buyers surveyed in intervention areas in Mumbai and Nagpur (60 percent in Mumbai and 81 percent in Nagpur) noted increased risk of getting caught by the police due to stricter law

²³ Measuring Deterrence for Commercial Sexual Exploitation of Children in Maharashtra. 2020. Funded by a grant from the U.S. Department of State and awarded by GFEMS.

enforcement in the form of raids on brothels, massage parlors, and hotels over the past year when compared to non-intervention area (52 percent, Pune).

Project reports note that, in a similar vein, Blue Dragon trained legal aid staff, lawyers, government officers, and social workers on trafficking issues, and collaborated with police to support rescue and repatriation activities.

Working directly with survivors, vulnerable populations, and their communities: At least two projects worked directly with survivors and at-risk communities to raise their awareness about sex trafficking issues. According to the India CSEC subawardee theory of change, as well as logframes and key performance indicators reported by the project, the India CSEC subawardee provided paralegal support to survivors, supported case reporting and follow-up, and engaged with local communities to build their awareness of trafficking risks and reporting mechanisms.

Project documents suggest that Blue Dragon, for its part, worked directly with vulnerable children and youth through school-based prevention interventions, provided vocational trainings and livelihoods support to communities, and trained legal aid staff, lawyers, teachers, and students on trafficking issues and mechanisms.

2.4 What are key strengths and challenges to PEMS 1 implementation to date?

Subawardees showcased five key strengths, which combined to support the quality of PEMS 1 implementation, even in the face of challenges like the COVID-19 pandemic and government policies and regulations.

4

All projects showed resourcefulness in adapting activities in response to challenges and opportunities encountered during implementation while remaining true to their theories of change.

Evidence from project reports and interviews with GFEMS and subawardee staff suggests that all projects adapted their activities throughout implementation, in different ways, and for different reasons. Project theories of change, modifications made to logframes, and quarterly reports demonstrate that adaptations occurred at the implementation level. Project models remained in place through the evaluation period. As an example, staff from the India CSEC subawardee noted that the project's goals and objectives never changed, despite the many modifications to activities that occurred throughout implementation:

"We were all very clear, what we were doing, why we were doing, where were we heading...[we] never changed the goal post. I think that is very crucial, the goal post has to remain the same."

EnCompass found that these adaptations were conducted out of need, using data when available, or creatively thinking outside of the box, and intentionally, following an adaptive management strategy. Finding 10 describes adaptive management practices as a key strength of projects and the GFEMS strategy. Finding 11 describes ways adaptive management was implemented through GFEMS' and subawardees' use of evaluation and learning research to provide timely data and inform decision-making.

5

According to subawardees, sub-subawardees, and GFEMS, relationships with and buy-in from key government partners facilitated project implementation.

Subawardees, GFEMS, and TIP Office respondents noted that securing the buy-in of government partners and maintaining strong relationships with those partners facilitated start-up and implementation for the Neev Consortium, Blue Dragon, and the India CSEC subawardee. Buy-in looked different in different contexts; for example, because of complementary support schemes provided by government agencies, or as a result of policy agendas in line with anti-trafficking aims.

The Neev Consortium partners cited close collaboration with state government authorities as a project strength, noting that collaboration:

"...helps a lot because as you know the current environment in our country, if you raise any kind of complicated issue like the slavery or bonded labor, it is very difficult to fight for that. But, if you collaborate very well with the state, then your challenges and problems reduce."

This assessment was not able to demonstrate, based on focus and timeline, the performance and outcomes of this collaboration.

Several staff from the Neev Consortium mentioned that government alignment with project activities and aims was an essential complement to the project's efforts to support microcontractors. Several government departments and governance units were described as especially helpful with trafficking victim removals:

"We have labor department, which is strong partner for Building and Other Construction Workers (BOCW) registration and legal support... If we talk about legal case, police department is strongly with us. If they are not there then we cannot rescue labor. We get their support from them; it is important for them to be aligned with our program. SDM [sub-divisional magistrate] and district magistrate, they are also aligned with our program for rescue. Social welfare department because all the schemes channelize through them and Panchayat ["panchayat" refers to a local governance unit] gets involved in it."

Similarly, India CSEC subawardee staff and subawardees highlighted that government support, expressed in various ways, facilitated project activities. According to one interviewee, the state of Maharashtra, for example, advocated strongly for the rights of children and women, and demonstrated victim friendliness on a regular basis. Further, legal authorities' mandates that all police stations, courts, and shelters should open legal aid clinics led to increases in survivors seeking aid, and state court decisions centering on the rights of children and potential victims encouraged the establishment of child-friendly courts and collaboration with Child Welfare Committees (CWCs), as noted in an FY2019 Q2 report. CWCs are quasi-judicial bodies functioning under India's Juvenile Justice Act 2015 that support children in need of care and protection:

"The government moved swiftly to give us official approvals on all three Child Welfare Committees (CWCs), thus putting us ahead of schedule. The Commissioner's Office has further directed all Deputy Divisional Commissioners of Pune, Nagpur, and Sangli to co-operate with [the India CSEC subawardee] to make the respective CWCs child friendly."

It is important to note that the India CSEC subawardee has been working in this region on these issues for over a decade. Data do not directly attribute these changes to PEMS 1, but previous work and resourcing from PEMS 1 led to this occurring during the PEMS 1 project period.

Finally, in Vietnam, Blue Dragon staff and project partners benefited from alignment between the project's activities and aims, and provincial-level government policies in Ha Giang, ensuring that the project could more effectively secure the permissions and support needed to implement activities. Blue Dragon staff reported that their ability to effectively partner and coordinate with and across multiple local agencies, partners and organizations, especially in terms of supporting/facilitating inter-agency coordination and collaboration, was both in part a factor of the organization's reputation and an important strength of the project. For example, according to one interviewee:

"The strongest point is that we work with all relevant stakeholders at the same time. For example, if we focus only on working with DOLISA and neglect the work with the police, there would be no victims for DOLISA to work with."

An important caveat to this finding is that while these data show that subawardees described relationships and buy-in from government to be a key facilitator to project implementation, this evaluation did not have data from these government agencies to confirm their role and level of buy-in. The ILO Consortium research, for example, showed some limits to government cooperation and willingness to engage in recruitment-related anti-trafficking efforts associated with overseas labor migration to Hong Kong. There are few CSOs permitted to operate in this space and it is not clear what limits these CSOs have before losing buy-in from government agencies.

6

Subawardees that enjoyed a positive reputation with government agencies and other NGOs built trust with key stakeholders and collaborated effectively, which ultimately strengthened project implementation. Engagement of for-profit and NGO organizations and the partnerships they can forge have expanded the base of private partners engaged in counter-TIP efforts.

Positive reputation with government agencies and other NGOs: When subawardees were viewed positively by other players in the anti-trafficking space, they appeared to develop trust with key partners and stakeholders, as well as with beneficiaries, to the benefit of their projects.

According to GFEMS staff who conducted research and used criteria (see Finding 1) to select subawardees, Neev Consortium members were already trusted partners in the construction sector. As a result, news of the project spread by word of mouth within targeted populations, as microcontractors let other workers know about the project and its services. This appears to have expanded the project's reach, according to several interviewees.

India CSEC subawardee stakeholders also report that reputation and relationships were essential for project implementation. For example, according to a project staff member, India CSEC subawardee and partners had strong enough relationships that they decided to invest in a participatory, consultative project development process. This process was stated by the India CSEC subawardee and GFEMS to have helped various project stakeholders to develop a shared sense of commitment, buy-in, and support for the initiative, which further strengthened their relationships.

In the Philippines, Blas Ople also appears to have benefited from a positive reputation, as well as an intentional approach to further strengthening relationships with other stakeholders. A GFEMS interviewee reported that Blas Ople:

"... is known for its field of work in the migrant sector. They are real advocates of the issue, whether or not GFEMS is here to fund them, they are a strong organization by itself."

Partnerships with IACAT, the department of Social Welfare and Development, and the Overseas Workers Welfare Administration helped Blas Ople more effectively support case coordination, and work toward desired goals. According to one interviewee:

"The partners of the Ople Center, they are happy that we have a good project. For instance, the government agencies, the IACAT, the DOJ, they are happy that the Ople Center has offered a good project to them like our project with GFEMS, the ICMS. This government agency that referred to us a case of an OFW which should be theirs in the first place, that's a big thing, that the case is referred to Ople Center, that Ople Center is more trusted in terms of providing assistance."

Positive reputation with the private sector: The Neev Consortium also benefited from its efforts to develop partnerships with the private sector and lending organizations over the course of the project. Doing so provided access to working capital to support microcontractors. As described in a quarterly report, the project leveraged its partnership efforts to establish relationships with:

"...multiple lending organizations that could provide access to working capital/liquidity support to needy MCs."

Blas Ople staff also highlight their reputation with private sector actors involved with anti-trafficking efforts.

"Our number of partners also increased, for instance in reintegration, we have big companies like San Miguel, so it's really a big help, we are already known and we are becoming more known."

Evidence provided by FEF and BITC/SHA indicates that positive reputations and intentional approaches to facilitating and sustaining cross-stakeholder collaboration strengthened project implementation. Additionally, the for-profit business engagement serves a valuable addition to partnering capacity for counter-TIP efforts. As a BITC/SHA project staff member mentioned:

"We are an organization who works not just with the not profit side...but also...with a lot of businesses... I think for a Nonprofit partner it is not easy to reach out to businesses and get them on board. But I think out here this was the unique role we were playing or this a forte that the alliance has, which has the option to bring these partners on board and also kind of it is not just for once, it is more like there is an opportunity and this is something which can be an opportunity for these businesses as well..."

7

Projects benefited from subawardees' sectoral and subject matter expertise, strong networks within sector-geographies, and contextual understanding.

Strong sectoral expertise, positive reputations in local contexts, and GFEMS advocated and supported adaptive management capacity on the part of subawardees contributed to the implementation of PEMS 1 subawardee projects.

Expertise, experience, and local networks: Several PEMS 1 subawardees had strong expertise and experience in their selected sector-geographies, including a knowledge of local contexts and established networks with others working on anti-trafficking initiatives in government and beyond.

The India CSEC subawardee and GFEMS staff explained that the India CSEC subawardee's in-depth knowledge, expertise, and experience in the human trafficking sector and its well-established, positive

working relationship and networks were crucial strengths, especially in collaborating with government agencies and other key stakeholders. In an interview, a project partner noted:

“[There was] no way otherwise we could have gotten the permissions and you know managed so much coordination in the field in a very, very, short timelines without the strength of these relationships...”

BITC/SHA-Mumbai’s long-term experience working with hotels and businesses, as well as with trafficking survivors and at-risk populations, reportedly strengthened the implementation of project activities. According to one GFEMS interviewee:

“The strength that I think is probably the most important to this even being a viable idea is their ability to engage with... the world’s leading hotel companies and get them on board with such an idea ... having them corporately commit to supporting survivors of trafficking was critical to be able to do this... So that’s one thing that I think was the foundation upon which this project could be built, was being able to bring in those top five-star hotels.”

BITC/SHA-Mumbai’s strong local networks and partnerships, including with project partners and sub-subawardees, also appears to have strengthened the project’s ability to identify and mobilize potential beneficiaries for project activities. One project partner noted that:

“Access [to beneficiaries] is easy through these partner NGOs. That is one major thing, and also, we know that ... an NGO like Prerana or [the India CSEC subawardee] is going to support [beneficiaries] given that they are the bonafide students of such organizations. That genuine thing also comes into play; we are not wasting any of our seats ...”

Contextual understanding: In a similar vein, GFEMS, TIP Office, and Blue Dragon staff also cited Blue Dragon’s in-depth, grassroots experience in the anti-trafficking space as a strength. According to a GFEMS staff member:

“I think that the greatest strength is the very strong local knowledge and context that Blue Dragon, its partners in Women’s Union and DOLISA, which is the Department of Labor, Invalids, and Social Affairs in Vietnam—that they really bring, both to obviously informing the project but well as in terms of informing the [Evaluation and Learning studies] data collection plan, informing the tools, the surveys, et cetera. I think that there was really a strong presence that they have on the ground in the communities that they work in.”

8

Domestic policies and regulations, some having to do with the COVID-19 pandemic, as well as lack of understanding by some subawardees of U.S. policies, affected the ability of most projects to carry out all activities as planned.

Several contextual challenges affected the implementation of PEMS 1 projects. Most notably, COVID-19 posed a serious challenge for several projects, as did government policy and regulations. As discussed in Finding 10, in many cases, subawardees responded to these challenges by adapting planned activities.

COVID-19-related challenges: The COVID-19 pandemic and associated lockdowns and social distancing restrictions affected the implementation of several PEMS 1 projects, and in some cases, may have limited the extent to which intended beneficiaries were able to access services and/or support provided by PEMS 1 subawardees. The effects of the pandemic, however, varied across contexts.

In Vietnam, quarterly reports from Blue Dragon note that lockdowns and social distancing restrictions implemented by the government in response to the COVID-19 pandemic delayed many Blue Dragon activities, including school-based programming. That said, delays, for the most part, do not appear to have resulted in the cancellation of activities. When the government relaxed restrictions, Blue Dragon was able to move ahead with implementation. As described by one staff member,

“February, March, April we had delays because we couldn’t travel to Ha Giang. And even staff could not travel within the province, at least for March. And the schools weren’t in, so especially the school activities were very delayed. We just now (September) really got it going with school activities. So mainly, it was just delayed. We haven’t had to significantly change any activities. We have been able to stick with the approach and the models...”

Other projects, including India CSEC and FEF, described similar delays to project activities as a result of government measures associated with the pandemic. In their case, however, restrictions appear to have affected the quality and reach of some planned activities, or even led to their cancellation. In the case of the India CSEC subawardee, social distancing restrictions reduced the extent to which the project was able to provide direct victim services in childcare institutions and shelters, and as a result, intake cases and client referrals may have dropped. Group KII respondents explain that victim services may have been particularly affected:

“... a lot of victim services was not individual one-to-one activities of putting something in a box; there were a lot of group activities also. All of that was made impossible. So, the nature, the quality, the frequency, the feasibility, the relevance, the reaching out was seriously affected until the unlocking process [related to Covid-19] started.”

According to FEF reports, COVID-19 lockdowns forced the cancellation of several project activities, including training for domestic workers and scheduled assessments, and also halted the development of the project’s Honest Jobs agency.

Domestic regulations and policies: Domestic regulations and policies affected several projects, delaying and/or hindering the implementation of key activities. In India, the India CSEC subawardee’s FY2020 Q4 report notes that the introduction of an advisory from the National Commission for the Protection of Child Rights posed challenges for project activities and beneficiaries, as described below:

“According to the advisory, all the children currently living in CCIs should be returned home to their families or registered for adoption or foster care after a due assessment within a period of 100 days. There is a risk of the advisory being misused leading to sudden restoration [to homes] adversely affecting the rehabilitation process for the victims of commercial sexual exploitation and sexual violence. The process may shift from direct social case work assistance for children in CCIs to supporting restoration of children to their home.”

The Neev Consortium also faced challenges related to government regulations. The local government in Delhi instituted a month-long pollution ban in 2019 that prevented construction activities, which caused workers to leave Delhi. This thinned the pool of beneficiaries with whom the project was working, and limited the project’s ability to reach as many microcontractors as planned. As described by one Jan Sahas staffer:

“Pollution ban was a complete ban on construction activities for at least a month. So, workers [had to] go back to their villages saying that they will come back once the ban is lifted. Once the pollution ban is lifted, getting all the workers back to the sites takes a lot of time, energy,

and money. So, if 100 workers have gone back to the villages then only 50 would come back. So that's a problem for MC as well as us."

In Vietnam, the BITC project was unable to recruit beneficiaries because of the federal government's decision to decentralize support for vocational training for trafficking victims in geographies outside of Hanoi. As discussed in previous findings, this led to a pivot from the closure of BITC Vietnam with funds reallocated to BITC India.

Lack of understanding by some subawardees regarding U.S. Leahy Law requirements: The

U.S. Government's Leahy vetting requirement and some subawardees' poor understanding of its likely impacts affected the implementation of several projects' enforcement related activities, including Blue Dragon's planned training for police. Leahy vetting is a process through which the U.S. Government vets foreign security forces nominated to receive assistance with funds appropriated in the Foreign Assistance Act, Arms Export and Control Act, or the National Defense Authorization Act to ensure they have not committed a gross violation of human rights. All recipients of U.S. Government funding must comply; many do, and most are able to carry out project activities. According to the TIP Office, they had conversations with GFEMS about this issue and expressed their concerns that organizations would not want to push the government to conduct this vetting.

As reported by Blue Dragon in a 2019 quarterly report, local officials refused to participate in the Leahy vetting process:

"The requirements to conduct Leahy vetting for each district police commander and individual commune police of 196 communes and all Border Guard officers of Ha Giang has been refused by Ha Giang government. That means Blue Dragon is not able to implement the training courses for Police and Border Guard officers."

Blas Ople's training faced delays while awaiting Leahy vetting results for police and investigators in the Philippines, as well as the India CSEC subawardee's Anti-Human Trafficking Unit trainings, though these activities eventually went ahead. Organizations more practiced in implementing U.S. policies, such as the India CSEC subawardee, could more ably navigate this issue, though differences in government responses to the Act may have also played a role. A GFEMS interviewee noted that:

"There's one component...Leahy vetting... that affected [the India CSEC subawardee] and really delayed the work. But unlike in the Blue Dragon case, we did eventually make it to a point where we could get the data that was required by the TIP Office and post in order to process the Leahy vetting. So that ends up in a happy story, but it did delay quite significantly the work in that area."

9

Subawardee and beneficiary respondents cited that several projects struggled to consistently communicate with, mobilize, and/or sustain the involvement of intended beneficiaries in vocational and job-skilling activities.

Interviews with subawardee staff and beneficiaries suggest that several projects struggled with engaging beneficiaries in vocational activities.

According to Neev Consortium microcontractors, for example, the project reportedly delivered some training programs using English language materials, which affected the ability of microcontractors to participate, and may have contributed to some microcontractors dropping out of the program. At the

same time, Neev Consortium subawardee staff noted that some microcontractors did not appear to commit to finishing training programs, in part because they might not trust staff members, at least initially. In an interview, field staff from Pratham describe the challenge in this way:

“This was the biggest challenge that for any individual to believe that we will not take any money from them and we will not make any demands to them... and we don’t expect anything apart from your time, these things are where we faced issues. The MCs are also busy as they always have jobs on their hands, running around for meetings, etc. so initially, we did face challenges to connect.”

Staff report that, over time, they overcame these trust issues.

Blue Dragon staff and beneficiaries also highlight language issues, and specifically, the lack of local-language translated communication materials, as a problem that restricted the degree of understanding by non-Vietnamese speakers of communications shared by trained social workers that were designed to raise awareness in communities about trafficking risks. As described by one staff member in a group KII, in one instance, Blue Dragon:

“... organized [a training] for 32 heads of the branches of women’s unions in villages and for members of the Women’s Unions in the three targeted communes. After the training, they conducted communications in their villages. However, many of them found it difficult to understand the training content because they could neither speak nor write Vietnamese language.”

Blas Ople project partners also cited limited commitment of some participants as a reason for dropouts, noting that, in a few instances, OFWs did not complete training courses because they wanted to return home. Finally, BITC in Vietnam, as previously discussed, was unable to recruit participants to its Hanoi-based training program, and eventually closed its Vietnam activities as a result.

10

Evidence shows subawardees’ willingness and capacity to adapt when faced with challenges, including GFEMS-supported use of adaptive management practices. These practices strengthened subawardees’ and sub-subawardees’ ability to implement project activities. Support and facilitation of adaptations by the awardee and the TIP Office played a critical role in providing space for adaptations.

Building on Finding 4, project subawardees and their partners demonstrated the ability to pivot in the face of external changes, and the ability to use evidence, mostly intentionally and systematically collected, to adapt project activities for success. This finding provides examples of subawardees’ willingness and gains in capacity in applying adaptive management practices and noticing its value, as well as adapting using other means. **Finding 11** provides specific examples of the use of studies, including evaluation and learning studies GFEMS developed intentionally to support adaptive management practices by subawardees, and ways in which using data from studies improved implementation.

Applying Adaptive Management Practices

What is adaptive management and why is it applicable for PEMS 1? When fully implemented, adaptive management is a structured, intentional, iterative process of robust decision-making for complex programs being undertaken in the face of uncertainty, with an aim to reducing uncertainty over time via

system monitoring. In this way, decision-making simultaneously meets one or more programmatic objectives and, either passively or actively, accrues information needed to improve future management.²⁴ Adaptive management is particularly appropriate for complex programs (systems) and those urgent enough that programming cannot wait until all uncertainty has vanished.

Adaptive management is a tool used not only to change a “system”—the program addressing needs and the context in which it is being implemented—but also to learn about the system. The use of theories of change for programs in context, providing a window into programmatic assumptions regarding how resources and activities in a particular context are likely to lead to short-, medium-, and long-term outcomes, can more appropriately inform evaluation and learning questions for project studies, and more easily allow for system understanding and modifications within the system, without moving away from planned objectives and goals.

Intentionally planned and iterative learning from ongoing monitoring, from selectively identified studies to address assumptions, and from interim evaluations of specific project components or approaches provides data that allows urgently needed programming to continue while learning is used to course correct or shift approaches in more timely ways. Because adaptive management is based on a learning process and includes an understanding of the system (described through the theory of change), it can address changes likely to improve longer term outcomes.

The challenge in using the adaptive management approach lies in finding the correct balance between gaining knowledge to improve management in the future and achieving the best short-term outcome based on current knowledge.²⁵ This approach has more recently been employed in implementing international development programs and has been adopted by USAID to include some, if not all, of its components.

Use of Adaptive Management in PEMS 1: Blas Ople, the India CSEC subawardee, BITC/ SHA, and GFEMS staff described adaptive management practices as significant strengths through which subawardees successfully responded to a variety of challenges. BITC/SHA and GFEMS staff described adaptive management practices as significant strengths through which subawardees successfully responded to a variety of challenges.

Blas Ople benefited from its adaptive management capacity by intentionally redesigning components of projects designed for survivor reintegration. According to a GFEMS interviewee,

“...adaptive management [was a strength] ... [For example], they had to basically then redesign the whole livelihoods component, and I think they did it really well and really thoughtfully and it’s become a really successful part of the project.”

Similarly, India CSEC subawardee project staff highlight the project’s flexible, consultative approach to project management as a strength that fostered collaboration across project participants, and enabled them to respond to a variety of challenges and changes throughout PEMS 1 implementation, as discussed previously in this section. A sub-subawardee associated with the India CSEC project, for example, described the project management approach as:

²⁴ Adapted from Holling, C.S. 1978. Adaptive Environmental Assessment and Management. John Wiley & Sons. ISBN 9781932846072.

²⁵ Adapted from Allan, Catherine; Stankey, George Henry. 2009. Adaptive Environmental Management: A Practitioner's Guide. Springer Science & Business Media.

“...very broad based, very participatory, very consultation based, very open kind of a thing...it was very clear, as to what’s the theory of change, of how do we want to achieve it. The logical conceptual structure was very visible and very transparent for anybody to see and question also, so that was something we definitely liked.”

The respondent went on to note that the project approach encouraged learning and accountability across all project stakeholders, saying:

“...that built a lot of accountability and made us think a lot. And also introducing us to...monitoring, evaluation, reviewing, learning, taking every kind of feedback from the beneficiaries. All of these things were great, learning opportunity for us...”

Facilitation of adaptive management practices by GFEMS and the TIP Office: While timelines and activities were managed as well as possible, as necessary adaptations were identified by subawardees, GFEMS and the TIP Office worked to provide no-cost extensions to subawardees requiring additional time to complete activities and achieve desired outcomes, evident from the numerous no-cost extensions granted over the course of the evaluation period. Additionally, GFEMS and the TIP Office accepted subawardee logframe changes needed to address necessary adaptations. This critical leverage provided to subawardees allowed the adaptive management environment to flourish. GFEMS’ facilitation of the adaptive management process is further discussed in Finding 11. As one India CSEC subawardee stakeholder notes, the positive working relationship between subawardees, GFEMS, and the TIP Office was seen as helpful:

“The entire project team, [the India CSEC subawardee] team, the GFEMS team, the people who are on board have actually been very supportive, very encouraging, and very understanding...”

Applying Other Forms of Adaptive Thinking

Below are examples demonstrating an adaptive culture that do not necessarily represent use of adaptive management practices, as these are reactive and responsive more than intentional and conducted through planned investigative and/ or experimental studies and planned decisionmaking methods.

Project modifications due to external factors: Some projects adapted due to external factors, such as the COVID-19 pandemic, Leahy vetting restrictions, poor understanding of donor rules and regulations, policy and regulatory compliance needs, and political unrest in targeted geographies.

For example, because of COVID-19, the Neev Consortium, the India CSEC subawardee, Blue Dragon, FEF, and Blas Ople delayed the timing of activities and pivoted to online, rather than in-person, activities. The Neev Consortium, the India CSEC subawardee, and Blue Dragon reallocated resources to provide emergency relief to those affected by the pandemic. See Finding 6 for more on how COVID-19 affected subawardee projects. Project reports and interviews also indicate that Blue Dragon and the India CSEC subawardee faced issues related to Leahy vetting restrictions. Blue Dragon adapted to these challenges by removing police training activities, whereas the India CSEC subawardee secured a no-cost extension to address delays. A GFEMS interviewee stated that when confronted with donor restrictions on providing direct cash assistance to support reintegration, the Blas Ople project instead decided to provide livelihoods training to OFWs.

An FEF quarterly report noted that when faced with delays resulting from Technical Education and Skills Development Authority (TESDA) regulations, the project adjusted the training schedule and operating

procedures to comply with new policies. An FEF staff member mentioned moving to more online efforts because of civil unrest in Hong Kong, which disrupted project activities.

Addressing recruitment: Several projects faced unexpected recruitment issues. For example, the BITC project in Vietnam concluded project activities because of persistent mobilization and recruitment challenges for its Hanoi-based training program. Survivors wished to stay near their families, which were not in Hanoi. These challenges arose prior to the pandemic and may have been further exacerbated by COVID-19 as the Vietnam government provided supports for families at their homesites outside of Hanoi. BITC's FY2020 Q4 report describes this adaptation and the reasons behind it:

"The project faced persistent challenges mobilizing and recruiting survivors for its training program in Hanoi. These challenges were compounded by a shift in VN government approach to providing support to trafficking survivors closer to their homes and in rural areas. As a result, shelters in Hanoi no longer housed survivors making it difficult for project implementation partners to reach these individuals... The project team concluded that the most effective way forward would be to conclude operations in Vietnam, and reallocate resources to focus on scaling up efforts in India."

This change in approach was generally seen as a positive by GFEMS and BITC, despite the issues it posed for the project. BITC reallocated funds and resources to the BITC/SHA project in India.

Other projects expanded the search for beneficiaries. The Neev Consortium expanded accessing recruits from destination (versus initially only from source) sites to increase recruitment. Quarterly reports and logframe changes show that the Neev Consortium identified additional migrant construction workers that the project might support, and built entitlement programming into destination work to reach more workers. BITC/SHA in Mumbai expanded its definition of target beneficiaries to include youth at risk of trafficking and trafficking survivors, and adapted an initiative to organize and coordinate with NGOs working with trafficking victims in India to reach more beneficiaries.

Addressing partnership issues: Subawardees encountered challenges working with project partners. In several cases, subawardees responded to these challenges by adapting planned activities and working out ways to adjust or change partnership plans in order to keep projects moving forward.

For example, interviews with Blas Ople staff and GFEMS suggest that, when faced with a lack of engagement and timely support from key partners, Blas Ople directly recruited and established new partnerships in order to find users for reintegration resources and ensure that activity timelines and targets were met. A Blas Ople staff member described this process at length in an interview:

"Initially when we designed the project, there was this intent to have recruitment agencies and non-government organizations come together and come up with systematic means to pull together their resources and come up with a systematic way of providing reintegration services to OFWs. However, as time went along, we had to adjust our strategies and we had to adjust our activities. We sort of went around CORE ... there's a time element and CORE wasn't really acting as fast."

Meanwhile, according to interviews, BITC/SHA adapted its programming to build more business partnerships and expand job placements beyond hotels, in order to ensure the usefulness of project activities and youth at-risk/ survivor job placements:

“The Alliance ...[thought] very creatively about where else, what other industries and sectors would there be transferrable skills so that we can have the project continue and especially those who had lost their jobs, who had already been placed in hotels... And...that’s something that I think they did really really well....”

2.5 What evaluation and learning questions and other research studies have been designed and implemented, and how are the results used?

*** Evaluation sub-questions reordered to support flow of findings**

All subawardees, in partnership with GFEMS, designed evaluation and learning studies, per the GFEMS strategy. In many cases, study designs evolved over time, including in response to the COVID-19 pandemic, and subawardees often appear to have used results to inform project adaptations.

11

Several subawardees, with GFEMS support and guidance, developed, conducted, and used the findings and/or results of evaluation and learning research studies to inform project interventions to: 1) identify and screen potential participants in project interventions, 2) generate actionable information on the effectiveness of project activities, or 3) work out whether and how to modify or adapt activities.

With adaptive management, it is important to continue with the work while also conducting studies to inform and improve efforts. Evaluation and learning studies were developed with this in mind. According to documents shared by GFEMS, each subawardee planned several evaluation and learning studies, 10 of which had been completed as of October 2020. These studies served several purposes, among them, to provide information useful for projects in designing, implementing, and adapting project components to achieve objectives. Subawardees appear to have used the results of completed evaluation and learning studies, especially for studies they were involved in designing and/or implementing. Annex 4 summarizes the evaluation and learning studies developed by subawardees, including methods, key research questions, and completion date.

Study-informed adaptations: According to GFEMS and Blue Dragon quarterly reports, as well as interviews with Blue Dragon, GFEMS, IST staff, and project partners, Blue Dragon used the findings of several research studies—including vulnerability surveys and assessments, the baseline study, and a household survey—to identify individuals and households vulnerable to trafficking, and connect them with relevant project interventions and support. These studies were intentional and designed to reduce uncertainties and improve programming through iterative learning, **per adaptive management practices**. Evidence can often not be corroborated, as study data were not available to EnCompass during the evaluation period. However, examples and volume of corroborated primary and documented data, along with a few studies having data and some beneficiary and partner input, show that the value of applying adaptive management practices to address project implementation challenges appears strong enough to support this finding.

Examples of adaptive management from Blue Dragon: Applying components of adaptive management, Blue Dragon applied survivor needs assessment results to develop a training program with modules for social workers of DoLISA and the Women’s Union of Ha Giang. Blue Dragon noted that social workers

reported increased understanding of the situation of survivors, the difficulties they face, and the importance of trauma-informed care after trainings.

Blue Dragon applied adaptive management practices to improve programming for community awareness raising about human trafficking. It used data from the **baseline household survey** to inform the selection of target villages for prioritized intervention, identifying 12 vulnerable communities that shared at least four of the following high-risk criteria: 1) low income, 2) at least one migrant per household, 3) migrants' passports taken, 4) migrants working more than eight hours, 5) not using placement agencies, 6) high percentage of households (25–79 percent) know someone who has been trafficked, and 7) not many government services available.

Blue Dragon used the results of the **Information Education Communication (IEC) evaluation** to assess the effectiveness of methods used, and to adapt methods for community education. For example, post-event surveys with more than 1,500 people provided data indicating that the project's initial approach for IEC community events was not very effective, and Blue Dragon shifted its communication approach to focus on a Women's Union unit-based approach in response to this feedback. The Blue Dragon team made revisions and adjustments to the content and format of community-level engagement and awareness interventions based on data collected through **ongoing monitoring activities**. These monitoring activities included pre and post assessments following awareness activities/events to collect feedback and assess knowledge acquisition. Specifically, results showed that larger-scale awareness events such as plays or music performances at community marketplaces that involved one-way communication between presenters and audiences were not effective at improving participant knowledge of trafficking risks or protective factors. Based on this, the project discontinued this type of community engagement to focus further on efforts through village-based anti-trafficking boards and school-based prevention.

A **mixed methods baseline survey in a planned treatment and control district of Ha Giang province and interviews in schools** showed Blue Dragon that video would be an effective medium for educating students about the risks of trafficking, and Blue Dragon began plans to collaborate with television departments to develop a video for schools to raise awareness and communicate about anti-trafficking in communities.

Examples of adaptive management from the India CSEC subawardee: the India CSEC subawardee quarterly reports and interviews and documentation from GFEMS describe the use of findings and lessons from a mixed methods assessment of the India CSEC subawardee's Rule of Law interventions, post-training surveys, community surveys, targeted gaps assessments, and a rapid COVID-19 assessment to inform and adapt activities and approaches.

The **India CSEC subawardee Buyers Arrest Research Report** brought out recommendations that the India CSEC subawardee acted on around the relevant laws (ITPA and POCSO) and their application to buyers and traffickers; specifically, the term "buyer," "customer," or "client" (or similar terms) of sexual services from a minor aren't terms mentioned so it has left some ambiguity. After completing the research for the buyer arrest project, the team created a list of experts on the subject with recommendations for pursuing arrests of customers who are commercially exploiting minors. The India CSEC subawardee held a **National Consultation on Creating Deterrence of Demand for Commercial Sexual Exploitation of Children** in Q1 2020 in Mumbai with more than 45 experts from across the country in attendance, including eminent members from the judiciary, social services sector, and academics. The India CSEC subawardee's buyers arrest research and consultation showed there is need for statutory clarity. The India CSEC subawardee has adapted to infuse this into its engagement with

stakeholders, though not strictly through training. Learning about lack of awareness of legislation applying to buyers led the India CSEC subawardee to incorporating information on arresting customers, along with the legal provisions for arresting customers during the training of the Thane AHTU police officers.

The India CSEC subawardee conducted **feedback calls with trained public prosecutors** that indicated that there was more reliance on on-camera hearings—which they advocated for as these are considered to be more victim-friendly—in the context of COVID-19. This information was incorporated into plans to further emphasize how to do this well in future public prosecutor engagements. The feedback also showed that in practice, video conferencing was not used in some cases due to lack of internet facilities in the courts, which showed a gap to be filled/solutions to be advocated for in order to fully reach the intended impact of victim-friendliness.

The India CSEC subawardee conducted **community feedback sessions with children served by CWCs** to understand what they would like to see in CWCs and incorporated that feedback into its plans for CWC training and infrastructure upgrades across the board. As noted earlier, CWCs are quasi-judicial bodies, under India's Juvenile Justice Act 2015. CWCs support children in need of care and protection.

The India CSEC subawardee collected **community feedback throughout its Turbhe (Mumbai suburb) community engagement/awareness raising work**. This work aimed to engage community members to identify and report CSE cases to police. The feedback showed that people continue to fear police due to negative past experiences. Therefore, reporting crimes to the police, even crimes against children, would be an unlikely course of action by community members. The India CSEC subawardee **researched options for the community to safely report crimes** without any direct contact with the police, realizing the need for a safe space within the system to report crimes anonymously, and after exploring it as an option, opened a Legal Aid Clinic in the Turbhe community for this reason. This turned out to be a success based on increased reporting of cases. The India CSEC subawardee continues to try to improve the community's lack of confidence in police, which it now understands requires more slow and steady interventions.

Examples of adaptive Management from the Neev Consortium: The NEEV project partners have utilized data from the **LMT study** on an ongoing basis to make a number of operational adaptations through the life of the project. Two examples of these include:

Registration of potential migrant workers at transport hubs and destination locations rather than source villages: Assessment of the India LMT data informed changes in the overall outreach strategy for the project. Instead of focusing on community mobilization at the village level, where individuals were often not likely to migrate for construction work in the short term, field teams shifted focus to transportation hubs (including bus stations and other prime transit locations) and labor *chowks* of destination cities to better target and identify migrant construction workers.

Delivery of entitlements and types of schemes: Analysis of LMT data revealed that a large proportion of workers were migrating with dependents (including children), leading to a project adaptation to deliver entitlement benefits at destination locations as well as at source districts prior to departure. The types of entitlement schemes were also expanded to include cash and cash-equivalent benefits that covered family members in addition to the workers themselves. Further, analysis of LMT data and a Jan Sahas monitoring system found that applications remaining open for more than three months were more likely to be rejected by the relevant government departments. Based on this, NEEV project teams further revised their priority list of entitlements to focus on those with minimal burden of

documentation for workers and with quick turnaround times for delivery of benefits. **Impact of subawardee engagement in research:** Subawardees' engagement in the research design and data collection process may also have influenced their use of research findings to inform interventions. Blue Dragon's baseline study, Social Worker Needs Assessment, and IEC post-surveys and evaluation as well as SHA/BITC's ongoing monitoring via their suite of surveys with beneficiaries and other stakeholders included involvement by subawardees in the design and implementation of the research activities, and subawardees used study results to adjust and target activities.

The Neev Consortium designed studies to inform project interventions. Project narratives report that the Neev Consortium intends to use the results of the LMT prevalence estimation study to compare forced labor prevalence across workers grouped into various nonrandom treatment groups. These results are expected to support decision-making regarding resource use for project activities.

Narrative reports and interviews with BITC/SHA staff and GFEMS highlighted that the YCI-Mumbai project used a suite of surveys developed with support from GFEMS and IST research (including starter surveys administered to trainees on enrollment; exit surveys conducted with trainees who have dropped out of the program; social worker observation forms; graduate surveys; follow-up surveys with graduates; and hotel partner feedback surveys) to provide ongoing feedback at every stage of the program from the survivors and at-risk trainees, and to track progress and measure longer term employment success. They also took on an approach of learning from "failure" by applying lessons learned from the cancelled YCI-Hanoi project to the YCI-Mumbai project. According to a BITC project staff member:

"We know that we have got a great example of success in India and we also got an example of failure in Vietnam. We are really looking at those two different angles and what works, and what has not worked is something that I am proud of. We are very much in the process of looking at that learning and it will emerge due course. But I think that is probably, for me at least it is one of the most exciting things about the project. It is a pilot project and there is so much that we have learnt and we can take on board for future programming that I think is really- really interesting."

The evaluation did uncover that YCI-Mumbai pivoted to expand criteria for recruits to increase participation in the program, and to partnering with the hospitality industries after COVID-19 removed hotel options for employment. It is not clear the extent to which learning from the Hanoi project influenced these decisions, nor other ways information has led to adaptations for this project, but this does provide a great opportunity to apply adaptive management practices to improve outcomes in India.

COVID-19 effects on evaluation and learning studies: While the COVID-19 pandemic contributed to delays and disruptions, Blue Dragon, BITC/SHA, and the [India CSEC subawardee](#) indicated that they were still able to conduct associated data collection and/or complete planned studies with only minor adaptations or adjustments. Adaptations and adjustments highlighted by subawardees included, but were not limited to, integrating COVID-19 safety measures into data collection processes and procedures and incorporating additional questions into research studies that specifically examined the impact of the pandemic on targeted populations. EnCompass obtained data from only two studies during the evaluation data collection period, however, limiting the team's ability to assess the use of data for decision-making.

GFEMS quarterly reports indicated that as part of their technical support for Blue Dragon, GFEMS and IST designed measures to reduce COVID-19 risk into endline data collection procedures. A GFEMS respondent also noted that they tried to build more qualitative methods into the baseline study, to better tease out the impact of COVID-19 on the program's implementation and results. Data from GFEMS quarterly reports and an IST staff member also indicated that GFEMS and IST supported BITC/SHA in incorporating additional questions in the follow-up surveys (administered on an ongoing basis to graduates of its training program) aimed at understanding how the COVID-19 pandemic and related restrictions in Maharashtra have affected their safety, health, and livelihoods. In addition, as part of the India CSEC subawardee project, sub-subawardee Prerana also initiated a rapid assessment study to help understand the impact of COVID-19 lockdowns on victims of CSEC and their families.

Effects of COVID-19 on some study results and utility: While data collection continued and some data were available for use by subawardees, one GFEMS respondent highlighted broader concerns around the impact of the COVID-19 pandemic in terms of the comparability of some of the evaluation and learning data and research. Specifically, the COVID-19 pandemic started during the middle of PEMS 1 project implementation and remains an ongoing issue in several of the projects' geographic context, such as India; in turn, comparing pre- and post-program data—for example, baselines, midlines, and endlines—may not be particularly relevant or appropriate in these contexts. Moreover, any data analysis or comparisons would also need to take into consideration the disruption, delays, and adjustments made to the implementation project activities; disruption of the projects' ongoing internal monitoring and data collection processes; and the broader impact of the pandemic in terms of exacerbating the already high levels of vulnerability of many of the project's target populations.

2.3 What projects or implementing organizations demonstrate potential to reduce trafficking, and how are they doing this?

**** Evaluation sub-questions reordered to support flow of findings***

As a process evaluation, the focus of the evaluation was on how the program was being implemented and was not intended to measure performance. Questions in interviews and reviews of documents did explore activities and project components showing a potential to reduce trafficking. Based on the data available during the evaluation period, limited data from beneficiaries and a lack of baseline and endline data, nothing conclusive can be said about early progress in reducing trafficking for any project. However, from this exploration, some data indicate that several subawardee project activities show mild or early signs of promise in reducing vulnerabilities of participants to trafficking. Also, as evidenced by project cancellations, not all projects and activities show signs of promise. Examples of both are described in the findings below.

12

Self-reports from GFEMS and project staff members across roles and partners indicate that direct support and services to survivors and vulnerable populations to address immediate and medium-term needs have potential to reduce vulnerabilities to trafficking for participants and targeted populations. A caveat: data is lacking to confirm this finding.

Staff and partners associated with Blue Dragon, the India CSEC subawardee, FEF, and Blas Ople highlight short-term direct support services, including livelihoods support, and community risk awareness-raising activities (Blue Dragon), vocational and educational support, psycho-social support, and family strengthening activities (the India CSEC subawardee), vocational trainings and trainings for domestic workers (FEF), and support services for reintegration of OFWs (Blas Ople), as having strong potential to reduce trafficking.

Staff from the Neev Consortium believe that direct support and engagement with survivors of trafficking or those highly vulnerable to trafficking, especially through entitlements and support to migrant construction workers and their families living in rural areas, appears to be helpful for immediately reducing vulnerability to trafficking:

“We do the work of providing government [entitlements] to [migrant construction laborers] and their families. So, their family members get the benefit of government schemes. If conditions of their house are not good then they will be helpless to do dangerous work. If we provide their families with entitlements, then to a certain extent they don’t do such work, they don’t take the option of becoming a bonded laborer.”

An LMT study was employed to assess the effectiveness of NEEV project interventions—in particular, entitlements, skills certification, and employment under ethical microcontractors—at reducing the risk of forced labor exposure. Labor outcomes for migrant workers whom these interventions engaged were analyzed against the broader participant pool in order to understand the extent to which the interventions had a protective effect. A final report of the India LMT study, which enrolled 92,846 migrant construction workers and successfully followed up with 17,788 of them to gather data on their migration and employment experiences, showed improvements for migrant workers engaged in interventions.

Through a **worker voice study embedded in the LMT study**, the research team was able to follow up with a sample of 2,103 workers who had been registered for social welfare schemes by a Jan Sahas worker voice study, 1,159 of whom confirmed actual receipt of benefits during the study period. Quantitative analysis of responses in comparison with a similar sample from the larger study cohort confirmed that receipt of social welfare benefits has a protective effect for migrant construction workers, reducing their vulnerability to forced labor. Results showed that workers who received entitlements were between 10 and 25 percent less likely to experience forced labor conditions than those who had been registered for, but had not received, any benefits.

The migrant worker study conducted **follow-up surveys with a sample of 269 workers** who received skills certification via recognition of prior learning (RPL) through Sambhav Foundation. Comparative analysis revealed that the prevalence of forced labor in this group was lower than in the non-intervention cohort, indicating that the validation of worker skill sets can help reduce their vulnerability to forced labor situations.

Staff and partners from Blue Dragon and Blas Ople noted that activities focused on broader capacity strengthening for other actors in the anti-human trafficking landscape—including employers, governments, and service providers—may contribute to these actors developing the skills and tools they need to sustainably take action against trafficking. GFEMS staff, for example, highlighted that Blas Ople’s efforts to build the capacity of the multi-sector task force to use the ICMS system may reduce trafficking in the longer term:

"It's going to vastly improve the way that they deal with trafficking victims on a systems level and really help the coordination between the different agencies but then also it's creating a holistic system so it's not so fragmented when a victim has to go through like, here's my government case and then here's reintegration so that we're treating them more as individuals rather than siloed trunks. So, I think that the system really has the opportunity to be scalable and certainly to be replicable in a variety of countries, not even just expanding its use throughout the Philippines."

Similarly, Blue Dragon staff and early study results show the implementation of early warning systems with anti-trafficking boards and within schools as promising. Finally, the India CSEC subawardee and sub-subawardee staff identified Child Welfare Committee interventions, including the training of CWC members and the government's ownership of the CWC initiative, as especially promising (see Finding 11).

13

Reports, early study data, and interview responses from GFEMS and subawardee staff members across roles, legal aid partners, the TIP Office, and subawardees, indicate that several PEMS 1 activities show potential for a) strengthening the capacity of criminal justice system actors to carry out human trafficking investigations and manage human trafficking cases, and b) increasing awareness and sensitization of criminal justice actors and communities to trafficking and supporting the use of more child and victim-friendly approaches.

Strengthening the capacity of the criminal justice system: While reducing the vulnerability of people who may be trafficked affects one part of the trafficking equation, increasing the risk of prosecution, conviction, and penalties to would-be traffickers and buyers is another. Several projects show mild evidence of bolstering the capacity of criminal justice system actors—police, prosecutors, judges, and legal aid providers—to support the identification and prosecution of traffickers.

A 2020 deterrence study funded through PEMS 1 for the India CSEC subawardee included 1,678 respondents (from interviews and surveys), as well as analysis of web and online data to uncover and identify supply and demand side trends regarding CSEC and the use of technology. Among other findings, the study showed that potential buyers surveyed in intervention areas in Mumbai and Nagpur noted increased risks of getting caught by the police due to stricter law enforcement in the form of raids on brothels, massage parlors, and hotels over that past year when compared to a non-intervention area (60 percent Mumbai; 81 percent Nagpur; 52 percent Pune). The India CSEC subawardee has been working in this area for over a decade, so it is not clear that PEMS 1-funded work was a contributor, but treatments were conducted through PEMS 1. An India CSEC subawardee staff member reported that as a result of PEMS 1 project successes, police were now far more interested in human trafficking, and even demanded more support and training from the India CSEC subawardee.

"The police are ready, they are interested, they are keen. They in-fact are pursuing us to continue provide training, so then they are asking for the curriculum..."

Blue Dragon's quarterly reports highlight that investigative support provided to police and the provisions of legal representation to victims may have contributed to successful prosecution and convictions. Project partners suggest that trainings for legal aid staff and lawyers improved the quality of victim representation. Respondents independently corroborated having brought survivors to the attention of

investigators and providing evidence in cases prosecuted since then, though the EnCompass team does not have independent confirmation of this.

GFEMS staff further report that they have preliminary evidence that use of the ICMS may reduce the amount of time it takes trafficking cases to make their way through the criminal justice system in the Philippines. EnCompass has not seen this data to confirm. Blas Ople staff observed that the combination of the task force and ICMS has improved the rapidity of sharing data and the extent to which government agencies identify cases for investigation, gather evidence, and track cases through the criminal justice process.

“There is a tracking system when cases are filed so you are able to monitor. For example, we would like to see what happened to case A, so it’s okay it’s already with the fiscal -- I’m just not sure if that’s the term for the criminal ... So really, monitoring could be conducted fast, unlike when there’s no ICMS, you have to go to – you have to call this agency, you will have to ask. In the ICMS, if you are allowed, if you allowed to request, you can monitor.”

Increasing awareness and sensitization within the criminal justice system: India CSEC subawardee staff and GFEMS respondents shared in interviews that the India CSEC subawardee, having been working in this area for more than a decade, may be improving the level of sensitization and victim friendliness within and across the justice system. In particular, public prosecutors, lawyers, and paralegals may now have stronger capacity to take up and deal with human trafficking cases, while the overarching justice system structure—including police, judges, and prosecutors—was described as more survivor-friendly and interested in investigating trafficking cases.

Blas Ople staff members see improvements, with guidance, in more sensitively and appropriately supporting child trafficking survivors through the investigation and prosecution process. According to one Blas Ople staff member:

“... the system is getting in place and [survivors are] being updated about their case from time to time. When I started working at the institution children were telling me that they do not know anything about what is going on in their case. Over time we have made systems, we know the people who write the information, we are building the capacities, and we are making sure that the probationer officer at the institution is conveying these things to children. We are making sure that the probation officer makes it a point to present these girls before the CWC from time to time. ... So, I think the system is getting strong in a way that they take things forward [even] if we are not there. So, I think a lot of small changes that I have seen, which I am saying small but have been very impactful. They have made a lot of difference for children and for the families who come and visit the institutions, and for all of them, I think anybody else can take it from here.” [survivors are] being updated about their case from time to time. When I started working at the institution children were telling me that they do not know anything about what is going on in their case. Over time we have made systems, we know the people who write the information, we are building the capacities, and we are making sure that the probationer officer at the institution is conveying these things to children. We are making sure that the probation officer makes it a point to present these girls before the CWC from time to time. ... So, I think the system is getting strong in a way that they take things forward [even] if we are not there. So, I think a lot of small changes that I have seen, which I am saying small but have been very impactful. They have made a lot of difference for children and for the families who come and visit the institutions, and for all of them, I think. Anybody else can take it from here.”

In Vietnam, Blue Dragon legal aid clinic staff reported that, as a result of project activities, legal institutions were now more aware of regulations on human trafficking, and that legal officers are more willing to coordinate, communicate, and take action on trafficking cases as a result of the project's work. While this information was not independently confirmed, according to one project partner:

"The awareness has changed. The legal aid officers now do their work with a more proactive attitude. In the simplest example, if the investigation agency has a case [that the legal aid officers should participate in], but did not contact them, they would call the investigation agency and ask 'why didn't you contact me?' ... If the investigation staff do not let the legal aid officers get involved, I will talk to the head of the investigation agency. The survivors have to benefit from their rights. You see, there are many changes."

Increasing awareness and sensitization within communities: An endline **quantitative household survey conducted in intervention district of Meo Vac and comparison district of Quan Ba in Ha Giang province** for the Blue Dragon project showed several key improvements.

Knowledge and awareness of trafficking risks and reporting mechanisms: At the community level, households had become more aware of trafficking and other risks associated with irregular migration abroad, including commercial sexual exploitation, fraud, labor deception, and forced marriage. Respondents in the intervention district were significantly less likely to say there were no dangers at all. Findings revealed a positive relationship between exposure to awareness activities and improved knowledge of trafficking risks and vulnerability drivers at project endline. The results also indicate that respondents grew more concerned about a range of trafficker tactics, including those that are not particularly common strategies in the region such as abduction and threats of violence. Additionally, the village-level early-warning system has been successful at informing help-seeking behavior in the target district: respondents at endline were more likely to be aware of whom to contact in case of a suspected instance of trafficking. Awareness of the provincial anti-trafficking hotline also grew dramatically, with 28 percent of respondents listing this as a reporting mechanism versus only 0.04 percent of respondents at baseline.

Risky migration: Findings from the household survey revealed that households in the intervention district of Meo Vac saw a substantial decrease in propensity for risky migration. Prior to the project, community members who migrated across the border in both target and comparison districts did so under relatively precarious conditions. At baseline, only 40 percent of households with a migrant family member in Meo Vac reported migrating for work with a contract compared with 77 percent of households in the comparison district, Quan Ba. Further, a large proportion of migrant households reported the confiscation of migrant IDs and documents (43 percent in Meo Vac and 39 percent in Quan Ba) and indebtedness prior to migration (17 percent and 10 percent of households, respectively). At endline, households in Meo Vac were significantly more likely to migrate with a contract (64 percent, a 24 percent increase), while Quan Ba registered a much smaller change over the same period (81 percent, only a 4 percent increase). Households with migrant family members in the project district also registered a decline in reported confiscation of IDs and were less likely to be indebted at endline.

While findings from this evaluative effort point to early indications of the effectiveness of certain project activities, the relatively short implementation period and uncertainties associated with effects of the COVID-19 pandemic mean that any observed positive changes cannot necessarily be attributed entirely to project activities.

GFEMS, subawardees, beneficiaries, and follow-up survey data indicate early signs that some vocational training, financial literacy, and soft skills training, along with facilitating stable employment, may have potential to make survivors and/or populations vulnerable to trafficking more secure, which may reduce the risk of re-trafficking.

Promise of targeted skills trainings: Vocational training, financial literacy training, and training focused on building soft skills, including confidence, appear to be contributing to improved security for some PEMS 1 beneficiaries.

Vocational training from BITC/SHA and the Neev Consortium may have improved beneficiaries' job prospects, for example. BITC/SHA staff noted that project participants regularly found and maintained employment up to six months after completing vocational and soft skills training:

"We have not had [an] episode of students going back into trafficking. So that is something that has not come forward, with this project. The last year's batch completed the program and were getting into jobs in the start of this year, when the entire pandemic happened. Things in terms of how they were planned had changed for everybody. We have not had any episode. We have had maybe few drop outs. But dropouts not because somebody wanted to get back into trafficking. So, that is something that has not happened."

Similarly, Neev microcontractors, in an FGD, stated that they received training on digital financial literacy, recordkeeping, and bookkeeping, and that the skills they acquired in these trainings helped them in their professional jobs. FEF beneficiaries also noted that project trainings improved their financial literacy, while FEF staff claimed that alumni tracked for six months following completion of the training decreased their personal debt upon program completion.

Staff from Neev Consortium subawardee Pratham reported that participating microcontractors enjoyed improvements in their working conditions, as well as more agency and security in their employment, stating in an interview that:

"They are careful and cautious now and know that this can put them at risk if they are taking a loan or something similar. They tell us that they are ok now and safe and getting timely payments, working at their own will, can take day offs when they want, work for 8 hours a day. So, in all, we see a positive response from them, and I feel as if they are safe now. And are aware of the risks if they give in under pressure or if someone takes their documents against their will."

The Neev migrant worker study collected **follow-up data for a sample of 525 workers employed with the microcontractors** trained by Pratham and Sambhav Foundation through the project. The analysis found that migrant workers who were employed under trained microcontractors faced lower forced labor risks than workers in non-intervention groups. These findings indicate that providing microcontractors with capacity-building support and training on fair labor practices can potentially generate non-exploitative employment environments for migrant workers.

FEF staff also noted that alumni of the project's job-skilling programs were between 15 and 25 percent less likely to be terminated from their jobs, though this evaluation cannot independently verify that figure. FEF conducted **regular check-ins on a quarterly basis with its deployed workers** through phone calls or Facebook Messenger as part of the ongoing monitoring for the project. These check-ins collected

information concerning the well-being of the worker and how they are getting on in their job, and provided workers the opportunity to ask questions and for FEF to understand improvements it could make to its services. The check-ins also allowed FEF to ascertain if workers had access to all their identity documents, and if they were safe and happy in their current job. Based on data from this internal monitoring process, the FEF team had no reports of forced labor or indicators of exploitation during the life of the project. Feedback surveys indicated a 99 percent satisfaction rate among workers who responded, and a 93 percent satisfaction rate from employers. FEF workers have also been enrolled as part of an ongoing larger migrant worker study (Philippines OFW LMT study) that is tracking the effectiveness of the FEF ethical recruitment channel and FTC training at limiting forced labor risks for OFWs. The study will be concluded at the end of 2021, after which more robust quantitative results will be available.

Soft skills training may also have contributed to beneficiary security by helping them understand their rights and how to advocate for and protect those rights. A domestic worker trained by FEF, for example, reported improved confidence and sense of dignity as a result of project trainings:

“Because they are telling us that you shouldn’t think or shouldn’t see yourself as just helper, but a worker that also needs to adjust, that they also need to adjust to you. That you need to be confident, so they don’t belittle you as a Filipino. Because there are many cases already that has happened to Filipinos where Filipinos were looked down upon which is not right. Because as a human, you should be treated like a human being. That’s what they inculcated in us at the training center.”

Microcontractors associated with the Neev Consortium reported that the project’s training and regular support, as well as helpline services and the Bandhu app, helped them secure employment, even amid the COVID-19 pandemic. A project staff member, in a KII, further noted that safety training covering safety gear, construction knowledge, and criminal justice procedure awareness and behavior trainings (soft skills) helped them secure and maintain stable, safe employment opportunities.

Promise of facilitating employment: Staff and beneficiaries associated with the Neev Consortium, BITC/SHA, and Blas Ople projects mentioned that project activities focused on facilitating employment opportunities—such as by linking jobseekers to potential employers or supporting survivors in starting their own businesses—may have helped at-risk populations secure and maintain more stable employment.

For example, the Neev Consortium worked with microcontractors and trained them on the use of an app, Bandhu, that helped connect microcontractors to potential employers, helping them secure stable employment opportunities and reduce their vulnerability to trafficking. According to respondents in an FGD:

“If there is any work for us, then we get a message. If it’s in our luck then we will get it but we will happily be looking at the opportunities. Earlier we had no information and the client didn’t know about us. The client did not have a lot of time and we would just look up on Just Dial and would directly call for work. So now we are closer to them and can get work faster.”

Follow-on survey data from the project suggest that most participants agreed that job placement activities were useful, and may have contributed to stable employment opportunities. BITC/SHA staff members in India report that, in their view, program participants improved their transferable skills and

job opportunities, including beyond hotels, and as a result, BITC/SHA alumni “are not going back into trafficking.”

A Blas Ople staff member mentioned that the project’s job-skilling programs appeared to contribute to beneficiaries’ efforts to develop their own businesses:

“70 percent of graduates have found either employment or alternative livelihood... 20 out of 25 graduates of livelihood and skills training program have established and have sustained their own businesses. But just take it with a grain of salt – it’s still early in the day, most of the challenges come later – it’s easy just to start.”

15

Several projects included livelihoods assistance components, but from the perspective of beneficiaries, the assistance provided was sometimes insufficient. Likewise, some vocational and skills training programs may not have fully met beneficiary needs.

Challenges for livelihoods components: Beneficiaries from the Blas Ople, Neev Consortium, and Blue Dragon projects noted that livelihoods assistance activities suffered from various flaws, from promised assistance not actually being delivered in some cases to assistance not being sufficiently comprehensive. The deficits in livelihoods activities rendered them less effective than participants would have liked.

Blas Ople offered several livelihoods training options to survivors, primarily in the food distribution and reselling area, or, in other cases, selling cleaning products. A Blas Ople beneficiary reported that the various cleaning products they were provided with to sell were not sufficiently in demand, and therefore were not useful, leading to the decision to migrate for labor after all:

“The products are still here with me. They were from a factory of dishwashing soap, fabric conditioner and detergent. However, it is not a stable source of income as there are also a lot of people selling these products. So, I have decided that if there is an opportunity, I will still go abroad.”

A GFEMS interviewee noted that the Neev Consortium’s program initially planned for microcontractors to be connected to working capital loans. Microcontractors associated with the Neev Consortium, participating in a group KII, noted that the assistance they had expected from the project did not materialize, despite their efforts to contact and engage the subawardee. According to a Jan Sahas staff member, this perceived gap may be the result of a misunderstanding about the support the project intended to provide:

“People also wanted certain financial support to sort of revive their business and that is not a part of the program. So that is something we could look at next phases of program.”

Others found that other supports were helpful. A Neev Consortium partner noted:

“The first point for me would be that the contractors are now working with proper documents and more professional work order so that payment is on time. In the case of workers, we have opened their bank accounts. They are taking money in their accounts and safe as cash could get misplaced. With bank accounts, they are planning for their future and their kids. So, this is a big achievement, not a complete one but slowly moving in that direction.”

Blue Dragon beneficiaries also described several deficits in livelihoods assistance. In particular, one group KII participant noted that though the project had provided livestock to some beneficiaries, no support for livestock medical care accompanied the animals. When livestock became ill, beneficiaries lacked the financial capacity to secure needed medicines.

Challenges with some vocational and skills training: Despite the success of some vocational and skills training programs, others appear to have suffered from a lack of language instructors and training not covering the critical topics needed to benefit from training or secure employment.

A BITC/SHA staff member noted that several beneficiaries complained that the project did not provide adequate computer skills training and that lack of expected direct cash assistance for trainings—unallowable per TIP Office policy that was consistently communicated to GFEMS—was problematic for them.

Blue Dragon staff and project partners explained that many potential beneficiaries did not appear to be interested in vocational training opportunities, and that dropout rates in training programs were quite high. Blue Dragon used data from post-training surveys to adjust its approach to focus more on local training centers and local employment, though EnCompass did not have data demonstrating this affected attrition.

An FEF beneficiary noted that the project prepared prospective domestic workers for the Overseas Workers Welfare Administration (OWWA) assessment, but included only basic counting skills in Cantonese. The lack of broader Cantonese language training posed challenges to passing the assessment for those seeking to locate to Hong Kong. According to one interviewee:

“...they should have taught us basic words like – the ones that were tackled in OWWA ... OWWA was very difficult, the instructors are very strict, your diction should be – very strict. And then when you make mistake, they punish you. That kind of strict – there are those who really fail at OWWA.”

Conclusions for PEMS Project Implementation

The conclusions below are drawn from across the findings presented in the previous section.

Conclusion 1. PEMS 1 subawardees applied a multifaceted approach to combating human trafficking, combining complementary, overlapping components with activities that engaged a diverse set of stakeholders at multiple levels. Although evidence is limited and the COVID-19 context is unusual, the approach shows potential for resilience in the face of challenges and some activities employed show mild potential for reducing vulnerabilities to trafficking for targeted populations and increasing risks to would-be traffickers.

Aligns to Findings 2, 3, 5, 12, 14, 15

Applying more complementary and comprehensive approaches across diverse stakeholders at multiple levels: Subawardee projects attempted to combat trafficking in persons using activities that aimed to reduce vulnerability to trafficking while also working to realign the incentives that lead to trafficking. Subawardees implemented this multi-track approach in different ways, corresponding to the sector-geographies in which they were working and their own skills and expertise.

While specific activities within subawardee projects were quite different from project to project (see project briefs, Exhibit 3 and Annex 3), in the aggregate, most PEMS 1 subawardee projects aligned with protection and prevention elements of the TIP Office's 3 Ps paradigm, along with developing partnerships. Projects included a mix of complementary activities, through which subawardees and their partners worked to increase short- and medium-term protections and reduce long-term vulnerabilities for trafficking victims. Projects worked directly with beneficiaries and with business, NGO, and/or government partners to address livelihoods and psychosocial supports and individual needs of survivors and vulnerable populations in their current state, and to provide work and career options employing ethical recruitment and labor conditions. A number of projects employed efforts to increase risks to traffickers by working to improve criminal justice practices (India CSEC subawardee, Blue Dragon) and reporting practices (Blas Ople), and to increase school and community member sensitivity to risks of trafficking (Blue Dragon). National and local government actors were engaged (Neev, the India CSEC subawardee, Blue Dragon, Blas Ople).

Projects' resilience in the face of challenges: PEMS 1 subawardees faced a variety of challenges during implementation. Unsurprisingly, the COVID-19 pandemic had the most significant impact on project implementation during this time period. The effects of the pandemic, and the restrictions, policy changes, and risks associated with it, depended in large part on the spread of the disease in particular contexts, and on the stringency of government restrictions imposed in response. The effects included delays to activities, shifts in targeted geographies, and reductions in the scope of some planned activities. The activities most affected were those most directed toward supporting people vulnerable to trafficking or survivors of trafficking, and many of the activities, such as providing assisted exit, psychosocial, and reintegration support (e.g., the India CSEC subawardee, Blas Ople, Neev Construction), were least amenable to moving to remote platforms.

Most projects made the necessary pivots by moving previously in-person activities to remote activities and using feedback from recruits, partners, and participants to ensure targeted individuals could and would participate. While delays might have affected some project components (e.g., Blue Dragon working with enforcement agencies, the India CSEC subawardee providing direct psychosocial care), other components for these more comprehensive designs (e.g., BITC/SHA's job-skilling, FEF's vocational training, Neev's livelihood supports, and various projects' evaluation and learning studies) could proceed as planned or with only short delays while pivoting to remote environments. Projects having less programmatic breadth, such as FEF, YCI-Hanoi, and YCI-Mumbai suffered strong effects on implementation, as these projects lacked other project components to continue when lockdowns reduced or eliminated existing partner options and participant options.

Blue Dragon, BITC/SHA, and the India CSEC subawardee also indicated that data collection and/or analysis for planned evaluation and learning studies continued during the pandemic with only minor adaptations or adjustments. This allowed, in some cases, for data supporting decision-making to be available earlier in implementation, allowing for earlier adaptations and course corrections for activities.

Early signs of potential success: Local and federal policies, along with U.S. policies and COVID-19, led to delays and additional challenges in bringing together vocational soft skills and criminal justice-related training efforts. However, training of prosecutors and enforcement agents, as well as increasing awareness and sensitivity of communities, showed early signs of promise in India through the establishment of more child and victim-friendly courts and following up on cases (India CSEC project), and interest in community engagement in legal proceedings (Blue Dragon). The development and training on the ICMS in the Philippines (Blas Ople project) and associated tools and protocols supported

more systematic and broadly used reporting systems and trafficking case management, as well as more rapid case management processes. An India CSEC subawardee CSEC study showed increased risks to would-be traffickers in treatment areas compared with nontreatment areas. Subawardees provide self-reported evidence of reduced rates of re-trafficking or being trafficked for survivor and vulnerable population alumni of the YCI/BITC project in Mumbai. Some Neev microcontractors reported improved profitability after engaging in training programs, improved soft and technical skills, and greater opportunity for gaining contracts that can employ more workers. Neev construction workers were reportedly able to make more thoughtful choices regarding employment as a result of receiving entitlements for their families, but this was not confirmed, as the team lacked access to construction worker participants.

Conclusion 2. Subawardees who lacked understanding of, or response to, U.S. policies faced challenges leading to delays, pivots, or lack of implementation of some activities; insufficient understanding of the local TIP context and survivor needs led to some project, activity, and beneficiary-level failures.

Aligns to Findings 2, 6, 11

U.S. policies: The TIP Office communicated U.S. policy requirements consistently with GFEMS. Subawardees lacking sufficient understanding or knowledge of U.S. requirements had more difficulty implementing and adjusting to these policies. Some subawardees lacked understanding of how to approach Leahy vetting or to propose activities that did not require this. Strong grounding in local contexts and U.S. policies allowed for some projects (India CSEC subawardee, FEF) to eventually carry out trainings with security forces. Blue Dragon was unable to secure permission for vetting and abandoned this activity. This led to a use of resources for an ultimately undelivered project component.

Some projects proposed activities that were not aligned to policies, such as direct cash transfers.

Failures and learning from failure: Programs were complex and conducted in challenging situations. Some also involved partners less grounded in the field of TIP (e.g., YCI/BITC, FEF, some Neev sub-subawardees). All projects also were faced with COVID-19-related challenges. Taking conditions and context as a whole, particularly for the groups newer to the TIP field, PEMS 1 projects represent a grand piloting of a strategy. While evaluating the TIP Office and GFEMS strategies for achieving change fell outside of the bounds of this evaluation, a recognition of this context is helpful in describing failures.

Approaches did not work in all cases. The YCI/BITC-Hanoi project is an example where lack of sufficient understanding of the Vietnam context for CSE survivors led to applying approaches and working with a subawardee unable to pivot sufficiently to match local conditions and survivor needs. A reallocation of resources provided an opportunity for this learning-focused subawardee to apply learning from one project failure to attempt to increase success for the YCI project in Mumbai.

In other cases, livelihoods training did not work for all participants (Blas Ople provision of goods to survivors to sell, but lack of markets for selling), and some supports for livelihoods options were insufficient to allow for success (Blue Dragon providing pigs for food or sale without sufficient veterinary care provision for participants). The assessment also uncovered a lack of understanding, research, or resourcefulness by some subawardees to address key language needs (e.g., Vietnamese overseas workers needing Cantonese diction training) and language barriers (e.g., Blue Dragon's provision of awareness-raising communications for communities in Vietnamese rather than the local language; use of English language training materials for Neev microcontractors). It is not clear if course corrections

have been made for these activities. These specific cases are noted among successes also noted by beneficiaries, but the beneficiary pool of respondents is too small to make more than thin conclusions regarding potential for success for project components (see Conclusion 1).

Conclusion 3. Five key subawardee strengths appear to have contributed to the implementation of PEMS 1 projects, including subawardees': 1) deep knowledge of local sector contexts; 2) emphasis on partnerships and local networks; 3) strong existing state and non-state relationships; 4) intentional approach to understanding, working with, and supporting key players in government and civil society; and 5) willingness and ability to manage adaptively when required.

Aligns to Findings 2, 5, 7, 8, 10, 12, 13

Where the evaluation was able to identify demonstrated early signs of progress and promise in PEMS 1 subawardees' work to address entrenched trafficking problems in complex contexts, those signs appear to be consistently associated with five strengths. These strengths supported implementation of activities when faced with challenges, as occurred for all projects.

First, subawardees were well grounded in local sector-geographies and/or the human trafficking landscape, demonstrating deep familiarity with the local contexts in which they were working. Second, subawardees consistently emphasized working collaboratively with local partners and networks throughout civil society, government, and the private sector, regularly leveraging these partnerships to tailor designs and plans to suit local needs and priorities. Third, subawardees enjoyed strong relationships with state and non-state actors *before* initiating PEMS 1 work, which they leveraged to quickly ramp up and involve others as projects kicked off. Early and ongoing buy-in obtained from a wide range of government and non-government stakeholders required strong extant reputations and networks by subawardees and sub-subawardees. Fourth, subawardees worked with their partners to intentionally identify and influence key players—in government, civil society, and the private sector—in local sector-geographies, carefully designing project activities to capitalize on and further strengthen existing political will for anti-trafficking initiatives. These five strengths appear to have provided a platform for a fifth, enabling subawardees to engage in adaptive management practices when required.

Conclusion 4. Subawardees, under GFEMS' strategy and guidance, incorporated adaptive designs and, in some cases, adaptive management practices. Subawardees utilized their knowledge, connections, and intentionally collected and analyzed data to adapt. Use of evidence-based adaptive management practices appear to have strengthened the implementation of subawardee activities.

Aligns to Findings 1, 4, 5, 7, 8, 10, 11

The focus of this evaluation was on process and implementation. The timing of the evaluation limited the data available to assess effects of implementation even on early outcomes. There is sufficient evidence, however, that adaptive thinking and adaptive management improved capacity for implementation.

As discussed in Conclusion 3, all subawardees faced challenges during project implementation. When encountering challenges, subawardees consistently adapted project activities, including by introducing COVID-19 protocols, finding new partners, and trying out new strategies for complying with policy regulations and engaging beneficiaries. Some of these adaptations were undertaken through creatively considering alternatives, such as expanding recruitment criteria and locales (Neev Consortium and YCI/BITC in Mumbai), and pivoting to accessible remote training methods during COVID-19 lockdowns.

In other cases, and as a key part of the GFEMS strategy for subawards, significant and intentional research, evaluations (e.g., post-training and/or follow-up surveys), monitoring (e.g., participant numbers, dropout rates), and addressing key learning questions co-created with subawardees provided data used by subawardees to inform decision-making more formally and systematically. This adaptive management approach required additional capacity and support by GFEMS, represented a capacity strengthening aspect of the project, was highly valued by some, if not all, subawardees, and was utilized by at least some subawardees to improve products, approaches, and delivery of services for beneficiaries.

Both means of adapting served to allow for shorter delays or more continuous programming and showed signs of making decisions based on changing contexts and listening to and learning from beneficiaries. PEMS 1 project evidence, primarily from subawardees themselves, but bolstered by several studies providing data and limited beneficiary and partner voices, suggests that successful adaptation does not emerge in a vacuum. Subawardee strengths discussed previously, including deep contextual and sectoral expertise, collaboration with networks and partners, and an intentional approach to collecting and using data to inform project implementation, contributed to subawardees' adaptive management capacity.

Importantly, it appears that the awardee, GFEMS, intentionally selected subawardees and projects that showed potential for learning as well as, in its estimation, effectiveness. GFEMS regularly encouraged and facilitated adaptive management practices through co-created and implemented learning and evaluation studies and expectations for course corrections. Also critical was the permission and freedom to adapt for program improvements granted by the TIP Office, leading to no-cost extensions, approved logframe changes, and recognition of the effects of COVID-19 on target outcome timelines. These facilitators and strengths led to GFEMS expanding reach and programming and increasing funding, including receiving additional funding from the TIP Office for PEMS 1.2.

An overlooked area, however, may have been knowledge or capacity-strengthening of newer subawardees and sub-subawardees to the TIP space regarding TIP context for their sector-geography. This lack of understanding appears to have led to some of the failures that appear to have been preventable (e.g., language-related training errors, lack of foresight about provisioning survivors with animal care training or support when providing livestock, or understanding market needs for food and cleaning products provided to or made by survivors to sell). Adaptive management practices aim to iteratively reduce uncertainties regarding programming within a particular context. Understanding of the context is critical to being able to undertake adaptive management and use data for appropriate decision-making.

RECOMMENDATIONS

Recommendations for Funding and Conducting Future PE Studies Based on the Evaluation of PEMS 1 Prevalence Estimation Studies

Evidence in this report demonstrates that designing and conducting PE studies in a TIP context is complex across TIP labor sectors, countries, and regions and is affected by policies, local social and cultural factors, languages and term interpretations, and the known aspects of collecting data to describe and understand often hidden and loosely networked populations. These recommendations translate the findings and conclusions identified from this evaluation into concrete actions researchers, practitioners, and funders can take to improve the appropriateness and quality of studies proposed and undertaken.

Recommendation 1: Prevalence estimation study implementers should include more thorough locally informed formative research (including gender and social inclusion assessments) and design pilot testing before undertaking a study, and the time and costs required for adaptations, to increase the likelihood of rigorous, representative estimates for the target of interest. Study funders should incorporate time and resources needed for locally informed formative research and design pilot testing into funding streams and RFPs, and researchers should include these plans explicitly in their proposals.

Aligned to PE Studies Conclusions 1, 2, 3, 4, 5, 6, 7

Formative research and pilot testing is a crucial step before rolling out a prevalence estimation study, especially as there are no simple answers to the best methodology for all contexts and sectors. The PEMS 1 researchers strengthened their study designs through formative research and pilot testing, but important gaps in formative research and pilot testing led to additional challenges in study implementation. Particular care should be taken to validate sampling frames and other key assumptions. Researchers should conduct gender assessments and analyses in the early stages of study design to help identify potential gender considerations needed for study rigor. Currently, gender analyses are listed as optional in TIP Office solicitations. The TIP Office should consider making these mandatory, especially for studies that are being conducted using novel methods or in new contexts. Formative research and pilot testing need to be tailored to the specific methodologies employed in the research.

Recommendation 2: Prevalence estimation study funders should ensure both experts in prevalence estimation and in human trafficking in the relevant labor industry and geograph are involved in prevalence estimation design and implementation. TIP experts should engage in discussions with researchers around challenges and decisionmaking needed to balance study breadth, rigor, contextual appropriateness, and representativeness to ensure funder objectives are met.

Aligned to PE Studies Conclusions 1,2, 3

As part of study design, funders need to determine the level of contextual relevance of the research and any trade-offs they might have for generalizability.²⁶ One of the elements of the EnCompass rubric for assessing prevalence estimation studies is that the planned use of the study is articulated before design, and that the design appropriately addresses that purpose. Funders and researchers should determine the impact they want funds to have. Studies need to be timed, scoped, and bounded appropriately to ensure key objectives are addressed and not diluted by other objectives. Because prevalence estimation research in the TIP context is complex, requiring numerous design decisions over time, continued engagement between funders, researchers, and external experts is essential to ensuring that design and implementation decisions continue to align with the study goals and planned use of the study findings.

Further, to achieve these goals, on-the-ground knowledge is critical to developing strong studies and cannot be replaced through other means. Funders should be aware that it may take more time up front to find and partner with effective local organizations, but doing so can be essential. In-country partners can be key to navigating contextual and logistical challenges, including local bureaucracies, understanding how and where people are moving, understanding issues about informal recruitment networks, how sex and other demographic characteristics may affect response rates, and other socio-cultural issues.

Recommendation 3: Prevalence estimation study funders and researchers should consider the tradeoffs between requests for prevalence estimation and other learning (e.g., intervention evaluations, real-time data, and trend data) to determine the possible costs and effects that integrating these will have on prevalence estimation resources and study rigor.

Aligned to PE Studies Conclusions 1, 2, 3, 4

PE studies require significant resources and time to conduct the necessary formative research, develop appropriate methods and sample, and, as shown here, when measuring TIP populations in most contexts, will face numerous recruitment, data collection, participant attrition, response bias, and analysis challenges, even with what appears to be a solid sampling frame. While collecting other data can also be useful, depending upon the type and amount, it can interfere with, or create additional challenges or costs to, ensuring rigor and representativeness.

If considering designs that support ongoing, real-time measurements of and interventions toward individual labor experiences, as well as prevalence estimation for a well-understood target population that is likely to sustain engagement, then a longitudinal study (heavily adapted to the local context and with an emphasis on research rigor) could provide opportunities for providing near real-time data and trend data on large-scale labor populations for a given industry and context while also supporting helpful interventions to improve the lives of individuals vulnerable to or experiencing modern slavery.

²⁶ This recommendation aligns with the following PEMFAC statements:

PEMFAC Statement GC16) Significant thought should be put into the intended purpose and use of prevalence estimation studies before they are conducted. The estimation method used should be guided by the purpose of the study and the level at which it is implemented.

PEMFAC Statement GC15) Research that highlights the prevalence of trafficking and forced labor and identifies modifiable determinants in specific sectors and geographies can help interventions target the most vulnerable populations and can guide private sector actors to better mitigate supply chain risks.

PEMFAC Statement GC21) Localized prevalence estimation research over time, if designed for the purposes of informing interventions, may provide relevant information on the effectiveness of programs. Localized studies can provide insight into specific social, cultural, political, and economic circumstances.

However, funders would need to be aware at the outset that the considerable challenges to rigor in the LMT methodology may mean accepting a less rigorous prevalence estimate as a cost for the other benefits of the method.

Recommendation 4: In the short term, prevalence estimation study funders should consider the potential gains of funding the same institution to conduct multiple studies with similar objectives, methods and contexts sequentially, as learning from one study can lead to efficiencies in another. For the long term, funders should require researchers to share practical lessons learned from study implementation widely so that other researchers and institutions can benefit from this knowledge.

Aligned to PE Studies Conclusions 1, 5

PEMS 1 research partners applied lessons learned from one study to the next when using similar methodologies in multiple studies. This is helpful in the short term for iteratively testing and refining methodologies to estimate the prevalence of TIP populations. In the long term, funders and practitioners in the field do not want this knowledge to be isolated at a limited number of institutions. Accordingly, researchers should document adaptations and transparently share learning with others in the field to continue to improve the use of prevalence estimation methodologies in the TIP context.

Recommendation 5: Funders interested in more rigorous, context appropriate, and useful prevalence estimation studies—for interventions or more broadly—should fund research to address numerous questions experts in this field seek to test and explore, while balancing the need for end-user engagement and ensuring research is useful for policy action. Funders should also support carefully planned information sharing across broader, more localized groups that can benefit from and use the information to reduce modern slavery.

Aligned to PE Studies Conclusions 6, 7

Funders should recognize that the methodologies are not yet fully developed for TIP contexts to get valid and reliable estimates and therefore should continue to fund research to build a better evidence base on what works for prevalence estimation in these contexts. Funders should continue to support the use of diverse and innovative methods for estimating prevalence of TIP populations and facilitate widely sharing lessons learned from these efforts. However, funders should balance efforts to refine methodologies and answer questions about methods that are transferrable between contexts with the needs of local stakeholders in specific study contexts by engaging with local stakeholders to understand what they need from research to support policy action. This may help funders increase the impact of research funds. As modern slavery is a sensitive topic in many regions of the world, information sharing requires careful planning with stakeholders to ensure this supports rather than hinders anti-trafficking efforts.

Recommendations for Funding and Conducting Future Projects to Combat TIP Based on the Evaluation of PEMS 1 Implementation and Early Progress

It is clear from the evidence considered in this report that human trafficking is a deeply complex issue, shaped by a variety of factors that vary over geographies and sectors and may shift over time, as situations evolve. These recommendations translate the findings and conclusions into concrete actions specific stakeholders can take to ensure effective programming and practices to address human trafficking.

Recommendation 1: Implementing organizations working to combat human trafficking should intentionally design for and incorporate adaptive management practices into projects, including short- and long-term evaluation and learning studies, ongoing context monitoring through local networks and partners, and regular reflection and learning moments. While we have limited data that this leads to more effective programs, other studies show this to be the case, and there is evidence that implementation of GFEMS' activities benefited from these practices.

Aligned to Project Implementation Conclusions 2, 3, and 4

Static anti-trafficking interventions, in which projects are designed and implemented without the flexibility to respond to new information and changing circumstances, are at best, unlikely to be effective, and at worst, could have a negative impact.

Data from this evaluation suggest that adaptive management practices contributed to more successful implementation of PEMS 1 subawardee projects, including by strengthening their ability to respond to unexpected challenges and to ensure approaches align to target population needs and styles. Successful adaptive management, however, does not happen in isolation. Rather, as evidence from PEMS 1 indicates, adaptive management appears to be undertaken and used when implementing organizations are grounded in local contexts, understand the role they play in the mix of partners and toward the larger project objectives, enjoy sectoral expertise, work closely with local partners to collect and use data to inform project design and implementation, and have grant flexibility to adapt.

Those working on TIP, therefore, should optimize their adaptive capacity by investing in building their local sectoral and geographical expertise, including by working hand in hand with local networks and partners, so that they are able to identify and respond to changes in context. They should intentionally design adaptive practices, including the regular collection, reflection on, and use of data, into their ways of working, to ensure that they are able to use evidence to inform decision-making processes throughout implementation.

Recommendation 2: The TIP Office should continue to work with implementers and other key stakeholders in ways that ensure funding and reporting timelines are adequately flexible, and implementers have the support, resources, and capacity they need to adapt to unanticipated challenges and new information, even if doing so means deviating from initial design plans.

Aligned to Conclusions 2, 3, and 4

All PEMS 1 projects encountered challenges. The COVID-19 pandemic, shifts in local and national government policies and preferences for PEMS 1 targeted countries and regions, and U.S. Government policies, including Leahy vetting requirements, affected the implementation of PEMS 1 projects. These kinds of challenges are not the exception, but rather, likely to be the norm, especially for projects seeking to address an issue as complex as human trafficking.

The TIP Office and other donors should therefore emphasize use of the adaptive approach reflected in PEMS 1, and ensure that funding agreements and modalities provide implementing organizations with the flexibility and learning opportunities they need to identify and respond to emerging challenges throughout project implementation. This might include funding longer term projects and supporting adaptive management approaches to MEL, as well as providing for resources and training to equip subawardees and their partners with the skills they need to design and implement fit-for-purpose learning efforts. These characteristics are, to some extent, already part of PEMS 1, but could nonetheless be further strengthened, built out, and integrated into subsequent funding rounds.

Recommendation 3: Funders and implementing organizations should carefully and intentionally: 1) resource and conduct formative research and relevant assessments, including gender and social inclusion assessments, 2) engage survivors and other local actors prior to and throughout the design and implementation of anti-trafficking initiatives, and 3) where safe and appropriate, cultivate partnerships across government counterparts, civil society organizations, and private sector actors.

Aligned to Conclusions 2 and 4

Funders should support, and implementers should design and carry out, key assessments and other formative research in advance of beginning projects to best understand target population needs in local contexts and, for migrants, for destination sites.

Value of understanding survivor needs, local TIP contexts, and cultivating key collaborations: Data showed that PEMS 1 projects with subawardees and/or partners deeply familiar with the TIP context for the sector-geographies in which they were working, and who emphasized partnership and collaboration with key players across government, civil society, and the private sector, could rapidly mobilize, gain partners, and gain trust among key stakeholder groups needed to drive change. These factors strengthened the breadth of the work they were able to undertake as well as the buy-in they could garner across diverse stakeholders. Robust local networks appeared to contribute to subawardees' positive reputations, and enhanced their capacity to gain new partners and their trust. Though data is limited, examples of beneficiary satisfaction appeared linked to the perceived alignment of training, support, and opportunities to specific beneficiary needs. The evaluation also yielded examples of beneficiary dissatisfaction that demonstrated a lack of this understanding. This speaks to the benefit of working with and building the capacity of awardees, subawardees, and contractors to have a strong, locally contextualized understanding of TIP for the sector of interest, and to cultivate strong local partnerships. The caveat to this is that anti-trafficking work can be sensitive and dangerous, so partnerships and collaborations need to be supportive and not create a more dangerous situation for implementers or targeted beneficiaries.

What implementers can do: As the key beneficiary population, where safe and of interest to them, survivors and representatives of those vulnerable to trafficking for sectors and geographies of interest should be consulted regarding designs and implementation approaches to anti-trafficking work. Implementing organizations should work to sensitively engage survivors and local actors in the design of anti-trafficking initiatives, and undertake participatory processes through which project stakeholders can collaboratively analyze implementation data, learn, and consider whether and how to adapt project activities over time. Implementers should conduct formative research, if not done already, including need assessments and gender and social inclusion assessments providing information about survivor and vulnerable population needs for the specific sector-geography, and of approaches and useful, safe potential partners for anti-trafficking work in this context.

What donors can do: Donors can encourage these practices by requiring local partnerships in grant selection criteria, and by ensuring that funded organizations have the space and flexibility to generate relevant pre-design data, learning opportunities, and design work, including, where safe and appropriate, alongside survivors and other local partners.

Recommendation 4: The TIP Office and other anti-trafficking funders should fund as well as build the capacity and interest of potential partners to undertake more comprehensive, responsive anti-trafficking initiatives in challenging contexts, as these designs show resilience in implementation when faced with external challenges.

Aligned to Conclusions 1, 3, and 4

Findings 2, 3, 5, 10, 12, 14, and 15, as discussed in Conclusions 1, 3, and 4, suggest five factors associated with the effectiveness of several PEMS 1 projects: contextual knowledge and expertise, partnerships and local networks, strong existing relationships with state and non-state actors, an intentional approach to understanding, working with, and supporting key players, and adaptive management capacity. These factors helped subawardees design and implement more comprehensive approaches to their anti-trafficking work, and identify and respond to emerging challenges throughout implementation, so they could sustain progress toward intended goals even as the operating environment shifted.

Not all organizations enjoy the highlighted strengths in equal measure. The TIP Office and other donors funding anti-trafficking work should consider ways in which they can ensure that selected organizations have the capacity to undertake more comprehensive designs. This might mean funding only organizations that already exhibit the five strengths discussed above. Or, more likely and more attractively, as it broadens and deepens the pool of organizations able to engage in complex and comprehensive counter-TIP efforts, it might mean that the TIP Office and other donors explore ways to further enhance the capacities of implementing organizations, including local actors, in these key areas. PEMS 1 has shown that newer actors not traditionally part of the TIP pool of INGOs, NGOs and CSOs, including private businesses, may be able to provide survivors with training, more ethical recruitment, and career opportunities that can reduce retrafficking. Engaging more businesses in this sphere can have a twofold effect of increasing the number of organizations engaged in the fight against human trafficking while removing those businesses from being complicit to or potentially engaging in trafficking practices.

In addition to seeking to assess whether candidate organizations already possess desired attributes, the TIP Office could provide additional resources for training and capacity-building as part of funding agreements. By further exploring modalities through which they, the funders, might support organizations on the frontlines of the struggle against human trafficking/modern slavery—directly or indirectly—funders could make further progress toward ensuring that their funds are going to the most impactful organizations and the ones best placed to address local problems and the drivers of vulnerability and human trafficking. If funders were to take such actions, they might also increase the number of such organizations, improve their sustainability, and strengthen the possibility that their work will make a difference.

CONSIDERATIONS FOR FURTHER WORK

Below are considerations for further work emerging from the evaluation team's work on the PEMS 1 process evaluation, but were not part of the evaluation itself. These represent inputs from the EnCompass team, including in-country consultants and Technical Advisory Group members.

Considerations for Funding Evaluation of Complex Efforts

Timelines and requirements for PE studies differ greatly from those needed for evaluation and learning studies. It would be useful to study the quality of PE studies and project work and outcomes when combining versus separating these two activities in terms of funding streams and contractors/awardees.

Combating modern slavery is a complex challenge that requires a coordinated, multipronged strategy. The TIP Office's approach to eliminating human trafficking involves the 3Ps, **p**rosecution, **p**rotection, and **p**revention, plus a fourth P, **p**artnerships. In the Notice of Funding Opportunity (NOFO) developed for PEMS 1, applicants were required to develop clear strategies to reduce the prevalence of modern slavery—including sex trafficking and forced labor of adults and children, men and women, and transgender individuals—in targeted populations. Strategies were expected to complement and build on existing anti-trafficking efforts, develop sustainable local capacity of governments and civil society, and incorporate survivor perspectives. The NOFO also required measurement of prevalence reduction. Applicants were required to present their plans for robust monitoring, evaluation, and reporting, and contribute data, analysis, lessons learned, and promising practices to support and amplify global efforts to combat modern slavery.

One Technical Advisory Group member noted that the TIP Office might consider developing two NOFOs for future PEMS evaluation work, one focused on prevalence estimation studies and a second one focused on anti-human trafficking programming. Experiences at the U.S. Department of Labor showed that these require very different skill sets, and EnCompass found that for PEMS 1, timing baseline PE studies work to coincide well with programmatic work proved challenging with both projects and prevalence estimation work underway. Having an additional period for targeted baseline prevalence estimation studies to occur prior to beginning programmatic work would support development of more effective baselines. This would also support independence of these estimates from programmatic efforts.

Results of this evaluation can also inform how Congress specifies the use of funds in order to give the TIP Office more flexibility in how to manage and distribute funds. For example, the TIP Office might address timing and commissioning of prevalence estimation studies separately from anti-trafficking project implementation.

Considerations for Utilizing TIP Definitions in Prevalence Estimation Studies

Projects and PE studies used a variety of techniques and criteria to include potential recruits based on interpretation and local applications of TIP definitions. An examination of TIP definitions used in PEMS 1 and other projects could support broader conversations about how to appropriately characterize and address TIP in local contexts while providing lessons and learning transferable to other funders, researchers, and practitioners.

Comments from the Technical Advisory Group noted that even at the time of recent legislation (e.g., 2020 National Defense Authorization Act), there was no official international definition of modern slavery, “and only recently have global estimates attempted to do so...but there are some thorny issues here with how the United States measures TIP, how NDAA defines it, and how the international community considers both TIP and modern slavery that have big implications for measurement.” Key informant interviews with researchers and project implementers identified numerous challenges in operationalizing definitions for measurement, particularly, balancing the need for appropriate contextualization of definitions to obtain accurate responses to survey questions, and having results or methods that are generalizable to other locales, populations, and contexts.

Considerations for Future Evaluation of Prevalence Estimation Studies

Fund evaluations to allow for greater breadth and depth of prevalence estimation study assessments.

Several suggestions regarding scope based on funding and time were raised during the proposal period, as we moved through the evaluation, and by the PEMS 1 Technical Advisory Group regarding scope for assessing prevalence estimation studies. These included:

- The TIP Office could consider commissioning research that was beyond the scope of the evaluation for more robust assessment of PEMS 1 prevalence estimation studies, including a deep dive into other prevalence estimation datasets from methodologically similar studies and contextually similar studies. The comparisons should include budget and costs, type of slavery, timeframe, and other considerations.
- Future research on prevalence estimation could also compare United Nations Office on Drugs and Crime (UNODC) or other databases of TIP cases to understand the relationship between prevalence estimations and cases that are documented in official systems. Furthermore, comparing these figures over time would also provide insights into the effectiveness efficacy of projects, policies, and the criminal justice system.
- Future efforts should include looking at rigorous and accurate ways to assess prevalence of modern slavery/TIP in different contexts, such as is being done with APRIES by comparing results of two methods for the same question and population.²⁷

²⁷ APRIES: African Programming and Research Initiative to End Slavery, University of Georgia (U.S.) and the University of Liverpool (U.K.)

Considerations for Evaluating Projects and Subawardees

Several areas of evaluation fell outside the scope of this evaluation or were not achievable within the context of this process evaluation, but would be important to include: a performance evaluation; a focus on the overall GFEMS strategy and ways it aligned to, diverted from, and expanded on TIP Office objectives; increasing use of survivor voice throughout the evaluation process; and additional forums for learning and sharing.

Conduct a performance and impact evaluation.

EnCompass is aware of and acknowledges the desire for performance evaluation-type data collection. This was a process evaluation that assessed the quality of prevalence estimation studies as well as progress toward project goals. EnCompass encourages a future performance or outcome evaluation of PEMS 1 or future PEMS, though, to avoid duplication, it should be noted that GFEMS may have its own evaluation processes. Auditing project-level M&E data would be one way to assess quality and accuracy of data systems and project implementation. Moreover, comparing baseline characteristics of the population and the beneficiaries/participants would also shed light on how well the project is serving the population and ensure that services are designed to address the needs and characteristics of the population of interest. Future work could also look at how intervention planning was informed by evidence-informed perspectives, such as the Minderoo Promising Practices database.

Include a focus on the overall PEMS 1 and GFEMS strategy.

EnCompass and the Technical Advisory Group recommend that future evaluations include examination of the broader PEMS 1 strategy and, specifically, GFEMS organizational strategy, since a project- and subawardee-level focus may miss key findings that could support future funding and management of these efforts, which show early signs of promise.

Insert more opportunities to include survivor voices in future evaluations.

Additional insights into the impact of PEMS funding would include understanding the characteristics and lived experiences of trafficking survivors supported by the program. This process evaluation hoped to collect some limited information from survivors and vulnerable project participants, but COVID-19 and concerns for survivor re-traumatization based on other evaluation work underway prevented these opportunities for most projects. Future work could also consider a broader array of stakeholders, such as state and local governmental agencies, funders, and CSOs to achieve PEMS' central goal of meaningful change. For example, government agency interviews could have helped corroborate information provided by GFEMS and subawardees, but because document review during the evaluation design period did not suggest that this would be important, government agency staff were not included in the interview process.

Use a variety of communicating events for learning, appropriate to different stakeholder groups.

Consider communicating and sharing findings with key external stakeholders: 1) project-level stakeholders who are only interested in their project and need a lot of detail and pragmatism in the findings in order to help them tailor future implementation or inform policymaking, and 2) the overall PEMS efficacy questions for the U.S. State Department and Congress. The types of presentations and materials will vary for these different audiences. All communication materials should accurately reflect the strengths and weaknesses of the evaluation design, but also condense the information in a way that is digestible for high-level stakeholders while still being meaningful.

ANNEXES

Annexes are provided under separate cover.

U.S. Department of State
2201 C Street NW
Washington, D.C. 20520